

**Synthesis and Evaluation of Designed (*S*)-Polycarbamate
Nucleic Acids, Erythrulose- based- and 2'-
Functionalized- Oligonucleotide Analogs for RNA
Targeting**

**Thesis Submitted to
The University of Pune for the degree of**

**Doctor of Philosophy
in
Chemistry**

BY

Venubabu Kotikam

Research Supervisor

Dr. Vaijayanti A. Kumar

**Division of Organic Chemistry
CSIR-National Chemical Laboratory
Pune-411008**

December 2013

CERTIFICATE

This is to certify that the work presented in the thesis entitled “**Synthesis and Evaluation of Designed (S)-Polycarbamate Nucleic Acids, Erythrulose- based- and 2’- Functionalized- Oligonucleotide Analogs for RNA Targeting.**” submitted by Mr. Venubabu Kotikam, was carried out by the candidate at the CSIR-National Chemical Laboratory Pune, under my supervision. Such materials as obtained from other sources have been duly acknowledged in the thesis.

Dr. Vaijayanti A. Kumar

(Research Supervisor)

Division of Organic Chemistry

CSIR-National Chemical Laboratory

Pune 411008

December 2013

CANDIDATE'S DECLARATION

I hereby declare that the thesis entitled “**Synthesis and Evaluation of Designed (S)- Polycarbamate Nucleic Acids, Erythrulose- based- and 2'- Functionalized- Oligonucleotide Analogs for RNA Targeting.**” submitted for the award of degree of *Doctor of Philosophy* in Chemistry to the University of Pune has not been submitted by me to any other university or institution. This work was carried out by me at the CSIR-National Chemical Laboratory, Pune, India. Such materials as obtained from other sources have been duly acknowledged in the thesis.

Venubabu Kotikam

December 2013

CSIR-National Chemical Laboratory
Pune- 411 008



Dedicated to
Family members and Friends

Acknowledgement

This thesis has been seen through to completion with the support and encouragement of numerous people including my well wishers, friends, and colleagues. At this point of accomplishment I would like to thank all those people who made this thesis possible and an unforgettable experience for me. It is a pleasant task to express my thanks to all those who contributed in many ways to the success of this study.

First and foremost, I would like to express my heartfelt and sincere gratitude to my research supervisor Dr. (Mrs.) Vaijayanti A. Kumar for all the advice, guidance, support and encouragement during every stage of this work. The confidence she had in me made to pursue the multiple roles of nucleic acids which always kept me energetic and interesting. Her energy, enthusiasm and scientific discussions are and would be the source of inspiration to realize my dream in this scientific world. I am highly thankful for her patience, motivation, timely advice and the continuous support provided during every stage of my research work. I could not have imagined having a better advisor and mentor for my doctoral study. Although this eulogy is insufficient, I preserve an everlasting gratitude for her.

Words fail me when I intend to express my thanks and appreciation to Dr. Moneesha Fernandes for her professional support and personal care. She was always beside me during the happy and hard moments, to push me and motivate me towards my goal. She was the one along with my mentor, from whom I learnt how to present a scientific idea and data. I believe the better way of thanking them would be through my future contribution to scientific community.

I take this opportunity to offer my sincere appreciation to Dr. Anil Kumar and Dr. (Mrs.) Vandana Pore for her help and valuable suggestions. I gratefully acknowledge Mrs. Anita Gunjal, and Mrs. Meenakshi Mane for their constant support, encouragement and especially for passing on their years of experience in handling the DNA synthesizer and HPLC analysis.

I extend my sincere thanks to the Director of CSIR-NCL Dr. Sourav Pal, Dr. Shivram (former director), Dr. K. N. Ganesh (Director, IISER Pune), Dr. Ganesh Pandey, Dr R. A. Joshi and Prof. D. D. Dhavale for their kind help and encouragement during the course of this work.

I owe a great deal of appreciation to our collaborators Prof. Michael Gait (MRC, Cambridge, UK) and Dr. Souvik Maiti (CSIR-IGIB, New Delhi, India) and their respective research groups for their exciting in vitro results of our antisense oligonucleotides.

My deepest gratitude to my teachers Prof. H. S. P. Rao, , Dr. C. R. Ramanathan, Dr. N. D. Reddy, Dr. Santosh Gharpure, Rama Sastri, Balakarun, Thomas for their inspirational teaching, guidance and blessings.

I also extend my sincere thanks to Dr. N. P. Argade, Dr. C. V. Ramana, Dr. B. L. V. Prasad, Dr. Hotha Srinivas, Dr. Sayam Sen Gupta, Dr. D. Srinivas reddy for their kind help and encouragement during the course of this work. I am also very much thankful to Dr. Mahesh Kulkarni and Mrs. Shanta Kumari and their group members for their excellent support in MALDI-TOF, HRMS and LC-MS analysis. The kind support from NMR group and the LC-MS facility, Organic Chemistry Division are also greatly acknowledged.

Research groups of Dr. Mahesh Kulkarni, Dr. B. L. V. Prasad, Dr. Manjusha Shelke and Dr. Vijay Mohanan are sincerely acknowledged for extending their instrumental facility and support during my new experience with graphene oxide studies.

My deepest gratitude to Dr. Moneesha Fernandes, Chaithanya Kiran, Dr. Seema Bagmare and Kiran Patil who has seen the beginning and completion of this thesis work and were even out of their comfort zone when I need a helping hand. My special thanks to Govind and Amit for their help during this thesis writing. I offer my sincere regards to people who have inspired me directly or indirectly in research carrier.

I have high regards for my seniors Dr. Khirud Gogoi, Dr. Madhuri, Dr. Sachin, Dr. Seema, Harshit and Namrata for their unconditional support and help during my Ph D. Course. I am indebted to my colleagues Kiran, Manoj, Anjan, Tanaya, Govind, Amit,

Harsha, Manisha and Ragini for providing a stimulating and fun filled environment. Also I would like to thank Dr. Raman, Dr. Amit Patwa, Dr. Ashwini, Dr. Sridhar Kosgi, Dr. Gitali, Dr. Mahesh, Dr. Roopa, Dr. Manaswani, Pradnya, Tanpreet, Deepak, Satish, Nitin and Vijay for their help and encouragement. I thank Bhumkar for the laboratory assistance.

My deepest gratitude to the families of Dr. Raman, Chaithanya Kiran, Dr. Madhuri, Dr. Seema, Dr. Rajender, Vilas, Dr. Santosh, Dr. Kosgi tambi, Dr. Raju, Dr. Ramesh, Namrata, Bhogesh, Suresh R R and friends Chaithanya Kiran, Kiran Patil, Ajay, Manoj, Priyadarshini, Ramakrishna M and Basavaiah Chandu for their love and caring.

I am indebted to my senior colleagues and friends Dr. Swaroop, Dr. Prasanna, Dr. Purude, Dr. Vidadala, Dr. Ramanujan, Dr. Ganesh Kokate, Yadagiri, Ravi, Narasimha Reddy, Ajay, Coach (Kanna), Suneel, Chandrababu, Srinu (gaali), Bala, Rami Reddy, Rambabu, Chaithanya Krishna, Jhony, Durga, Shiva, Dev Dutta, Nagendra, Narendra, Laxmi Prasad, Trinad, Bhaskar, Srinivas, Shrikant, Innaiah, Naresh, Dr. Sudhakar, Santi, Hanuman, Venkat, Vasu, Ashok, Satish, Kumar raja, Seetaram, Kasinath, Ankush, Dr. Prasad Wakchure, Dr. Abasaheb, Dr. Nagesh Kupse, Suresh K K, Sandeep, Swami, Kedar, Girish, Prabhakar, Pankaj, Priyadarshini, Saikat, Dayanand, Prakash Chavan, Sujit Pal, Kishanu, Shyam, Pravat, Mangesh, Pradeep, Chinmoy. My special thanks to Chaithanya Kiran, Coach, Suresh R, Trinad, Tejas, Komal, Pooja and Nishita for making my last year stay at NCL still memorable.

I extend huge thanks to my friends during school, College and Master's days.

I take this opportunity to offer my sincere appreciation to NCL Sports Club and the teams of Basket ball and Volley ball. I thank my teammates Alson, Pushkar, Kunal, Jogdand, P. D. Siddeswar, Krishna, Balraj, Giri, S. C. Mishra, Vilas, Chaithanya Kiran, and Kamteji for their support and team spirit during the tournaments. I also would like to extend my sincere appreciation to '*Deccan Dischargers*' players for their spirit.

I extend my sincere thanks to Student Academic Office and Division of Organic Chemistry at CSIR-NCL for their help and support at crucial times. I am also thankful

to the University of Pune for the smooth documentation process during my doctoral studies. I am grateful to CSIR, New Delhi, for awarding the research fellowship and Director, CSIR-National Chemical Laboratory for extending all infrastructural facilities.

Last but not least, I would like to pay high regards to my parents, Shri Venkateswara Rao and Smt. Jayalakshmi for their unconditional love, sacrifice, support and encouragement without which I never would be what I am. I also extend high regards to my brother Sadasivarao and sister Anusha and brother-in-law Muralidhar for their constant love, support and encouragement. Special thanks and wishes to my niece 'Hemasri Lakshmi' and nephew 'Santosh' who made my short visit to home entire memorable for the rest of the year during my doctoral studies. I owe everything to my family members.

Above all, I owe it all to Almighty God for granting me the wisdom, health and strength to undertake interdisciplinary research work for my thesis and enabling me to its completion.

Venubabu Kotikam

Contents		
Publication /Symposia		i
Abbreviations		ii
Abstract		v
Chapter 1		
1	Introduction to nucleic acids	
1A	Nucleic Acids: An Introduction	1
1A.1	Chemical components and primary structure of nucleic acids	1
1A.2	Base Pairing via Hydrogen bonding	2
1A.3	Sugar puckering in nucleos(t)ides	4
1A.4	Structures of Nucleic acids	8
1A.4.1	The duplex structure	8
1A.4.2	Triplex-Forming Oligonucleotides (TFO)	10
1A.4.3	Quadruplex structure – The ‘G-quadruplex’	12
1A.5	Spectroscopic technique to study DNA/RNA interactions	13
1A.5.1	Ultraviolet spectroscopy	13
1A.5.2	Circular Dichroism	15
1B	RNA therapeutics through antisense mechanism	15
1B.1	Introduction	15
1B.2	RNase-H activation	17
1B.3	Steric-blocking oligonucleotides	18
1B.4	RNA interference (RNAi)	19
1B.4.1	siRNA (small interfering RNA)	19
1B.4.2	miRNA (microRNA)	20
1B.4.3	piRNA (Piwi-interacting RNA)	21
1B.5	Splice-switching oligonucleotides	21
1B.6	Alternative splicing	21
1B.7	Some promising chemical modifications of nucleic acids	23
1C	Biomedical applications of synthetic DNA together with graphene/graphene oxide	27
1C.1	Introduction	27
1C.2	GO FRET Biosensor	28
1D	Origin of DNA/RNA	30
1E	The present work	33
1F	References	35

Chapter 2		
2	Applications of synthetic DNA in therapeutics	
	SECTION A	
2A	Design, synthesis and biophysical evaluation of 2'-O-(R/S-serinol) functionalized oligonucleotides	45
2A.1	Introduction	45
2A.2	Our design and rationale	47
2A.3	An efficient synthesis of 2'-O-functionalized pyrimidine nucleosides	48
2A.4	Synthesis of enantiomeric tethers from L-serine	51
2A.5	Synthesis of R/S-AMP phosphoramidite derivatives of uridine	52
2A.6	Synthesis of 2'-O-MOE and 2'-O-MOP controls	53
2A.7	Solid phase synthesis of oligonucleotide using phosphoramidite chemistry	53
2A.8	Synthesis of modified oligonucleotides, characterization and UV- T_m studies	55
2A.9	SVPD digestion of the modified oligonucleotides	58
2A.10	Conclusion	60
2A.11	Experimental Section	61
2A.12	Appendix	69
	SECTION B	
2B	Design and synthesis of 3'/5'-ends serinol capped oligonucleotides and 2'-(ω-O-methylserinol) guarded OMe-RNA/DNA mixmers for efficient splice correction	97
2B.1	Introduction	97
2B.2	A simple design and rationale	100
2B.3	Synthesis of amidite derivatives and modified oligonucleotides	101
2B.4	UV- T_m studies and SVPD digestion of capped oligonucleotides	103
2B.5	Splice correction activity of 3',5'-serinol and 2'-(ω -O-methylserinol) guarded 2'-OMe-mixmers	106
2B.6	Conclusion	108
2B.7	Experimental Section	110
2B.8	Appendix	118

SECTION C		
2C	Design, synthesis and biophysical evaluation of 2'-guanidino functionalized oligonucleotides	134
2C.1	Introduction	134
2C.2	Objective and rationale behind the design	136
2C.3	Synthesis of 2'-amino-2'-deoxyuridine	137
2C.4	Guanidinylation of 2'-deoxy-2'-aminouridine	138
2C.5	Significant conformational change- ¹ H NMR analysis	140
2C.6	<i>N</i> -Cbz deprotection under basic conditions and synthesis of phosphoramidite	141
2C.7	Synthesis of modified oligonucleotides, characterization and UV- <i>T_m</i> studies	143
2C.8	Conclusion	146
2C.9	Experimental Section	147
2C.10	Appendix	153
2D	References	169
Chapter 3		
3	Applications of modified synthetic DNA mimics in diagnostics	
SECTION A		
3A	Design and synthesis of (<i>S</i>)-polycarbamate nucleic acids ((<i>S</i>)-PCNA) and their biophysical evaluation	175
3A.1	Introduction	175
3A.2	Rationale and objectives of the present work	177
3A.3	Retrosynthetic analysis	178
3A.4	Synthesis of PCNA oligomers	179
3A.4.1	Synthesis of activated carbonate monomers	179
3A.4.2	Synthesis of <i>R</i> -PCNA thymine monomer	182
3A.4.3	Activated carbonates to carbamate backbone nucleic acids	183
3A.5	General principles of solid phase synthesis	184
3A.6	Solid phase synthesis of oligocarbamates	185

3A.7	Purification and characterization of the <i>S</i> -PCNA oligomers	186
3A.8	Biophysical studies of <i>S</i> -PCNA oligomers	187
3A.8.1	Binding Stoichiometry: UV Job's plot	187
3A.8.2	UV- T_m - Results and discussion	188
3A.8.3	Gel retardation of T ₈ complexes	191
3A.9	Conclusion	192
3A.10	Experimental Section	193
3A.11	Appendix	202
SECTION B		
3B	Evaluation of different backboned- and 2'-sugar-modified nucleic acids for DNA diagnostics in conjunction with GO	221
3B.1	Introduction	221
3B.2	Rationale and objectives of the present work	224
3B.3	Synthesis of the materials used	226
3B.3.1	Synthesis of graphene oxide (GO)	226
3B.3.2	Synthesis of labelled DNA mimic/analogue probe and MALDI-TOF characterization	228
3B.4	The study of short and different backboned-T ₈ -CF sequences towards probes development	230
3B.4.1	Studying the interactions of T ₈ -CF sequences with the GO using CD	231
3B.4.2	The fluorescence quenching and restoration of T ₈ -CF probes	233
3B.4.3	Specificity of the fluorescence restoration	235
3B.5	Studies extended to mixed pyrimidine/purine oligomeric probes	237
3B.6	Thermal desorption studies of the probes from GO surface	239
3B.7	Evaluation of longer DNA/DNA modified probes	241
3B.8	Conclusion	244
3B.9	Experimental Section	245
3B.10	Appendix	248
3C	References	258

Chapter 4

Design and synthesis of erythrulose- based thymine and adenine monomers, and the attempted synthesis of modified oligonucleotides

4.1	Introduction	263
4.2	Design of erythrulose-based nucleic acids and rationale	266
4.3	Retrosynthetic analysis-I	268
4.4	Retrosynthetic analysis-II	269
4.5	Synthesis of erythrulose- based monomers	270
4.6	Attempted synthesis of oligonucleotides	272
4.7	Conclusion	273
4.8	Experimental Section	274
4.9	Appendix	283
4.10	References	304

List of Research Publications:

1. “Robust synthesis of enantiopure cyclohexenyl analogues of 2/3-deoxyribose sugars as carbocyclic nucleoside precursors”, Manojkumar Varada, **Venubabu Kotikam** and Vaijayanti A. Kumar*, *Tetrahedron* **2011**, *67*, 5744-5749.
2. “Comparing the interactions of DNA, PNA and poly-carbamate nucleic acid (PCNA) oligomers with dispersed graphene oxide (GO)”, **Venubabu Kotikam**, Moneesha Fernandes and Vaijayanti A. Kumar*, *Phys. Chem. Chem. Phys.*, **2012**, *14*, 15003–15006.
3. “Synthesis and properties of 2'-O-[R- and S-(2-amino-3-methoxy)propyl] (R-AMP and S-AMP) nucleic acids”, **Venubabu Kotikam** and Vaijayanti A. Kumar*, *Tetrahedron* **2013**, *69*, 6404-6408.
4. “Enhanced splice correction by 3', 5'-serinol and 2'-(ω-O-methylserinol) guarded OMe-RNA/DNA mixmers in cells”, **Venubabu Kotikam**, Andrey A Arzumanov, Michael J Gait*, and Vaijayanti A Kumar*, *Artificial DNA: PNA & XNA* **2013**, DOI:org/10.4161/adna.27279.
5. “A novel 2'-Guanidino-functionalized oligonucleotides”, **Venubabu Kotikam** and Vaijayanti A. Kumar*, Manuscript under preparation.
6. “An efficient synthesis for 2'-O-functionalization of pyrimidine nucleosides”, **Venubabu Kotikam** and Vaijayanti A. Kumar*, Manuscript under preparation.

Patent Applications:

1. Locked and unlocked 2'-O-phosphoramidite nucleosides, process of preparation thereof and oligomers comprising the nucleosides. Vaijayanti A. Kumar, Anita D. Gunjal, Moneesha D'Costa, Namrata D. Erande, **Venubabu Kotikam** (2011): US Patent application 20110196141 A1.
2. 2'/3'/5'-Serinol functionalized nucleic acids. **Venubabu Kotikam** and Vaijayanti A. Kumar, US Patent Application filed.

Symposia Attended/ Poster Presentations:

1. Presented poster on “Efficient synthesis of enantiopure cyclohexenyl surrogates of 2'/3'-deoxyribose sugars” during **National Science Day celebrations** at CSIR-National Chemical Laboratory, Pune, India, February 2011.

-
2. Participated in "**1st CRSI zonal meeting**" held at CSIR-National Chemical Laboratory, Pune, India, May 2011.
 3. Presented poster on "Preliminary evaluation of nucleic acid oligomers containing unnatural backbone for DNA diagnostics using GO" during **National Science Day celebrations** at CSIR-National Chemical Laboratory, Pune, India, February 2012.
 4. Presented poster on "Nuclease resistant third generation modifications in DNA/RNA" during **National Science Day celebrations** at CSIR-National Chemical Laboratory, Pune, India, February 2012.
 5. Presented poster on "2'-guanidino-functionalized oligonucleotides" in **XX International Roundtable on Nucleosides Nucleotides and Nucleic Acids**, Montreal, Canada, August 2012.
 6. Attended "**ACS on campus**" meeting held at CSIR-NCL, Pune, India, October 2012.
 7. Presented poster on "Designed Synergy of Chirality, Charge and Hydration Modules: Synthesis of 2'-O-[R/S-(2-amino-3-methoxy)propyl] (R-AMP and S-AMP) Nucleic Acids" in **19th ISCB International Conference (ISCBC-2013)**, Udaipur, India, March 2013.
 8. Presented poster on "Designed Synergy of Chirality, Charge and Hydration Modules: Synthesis of 2'-O-[R/S-(2-amino-3-methoxy)propyl] (R-AMP and S-AMP) Nucleic Acids" in **International Meeting on Chemical Biology** at IISER Pune, India, May 2013.

Abbreviations

μL	Microliter
μM	Micromolar
A	Absorbance
Å	Angstrom
A	Adenine
Ac	Acetate
ACN	Acetonitrile
AMP	2-Amino-3-methoxypropyl
ANA	Arabinose nucleic acids
AONs	Antisense oligonucleotides
ASO	Antisense oligonucleotide
ATP	Adenosine triphosphate
Bn	Benzyl
Boc	<i>tert</i> -Butyloxycarbonyl
bp	Base pair
BPB	Bromophenol blue
BSA	Benzenesulfonic acid
C	Cytosine
Calcd.	Calculated
cMOE	Constrained-MOE
cDNA	Complementary DNA
CNTs	Carbon nanotubes
CPG	Controlled pore glass
Cbz	Carbonyloxybenzyl
CD	Circular Dichroism
CF	5,(6)-carboxyfluorescein
Conc.	Concentrated
dA	Deoxyadenosine
dT	Deoxythymidine
DCM	Dichloromethane
DEPT	Distortionless enhancement by polarization transfer
DI	De-ionized
DIPCDI	N,N'-Diisopropylcarbodiimide
DIPEA/DIEA	N, N-Diisopropylethylamine
DMAP	4-(N,N-Dimethylamino)pyridine
DMF	N,N-dimethylformamide

DMD	Duchenne muscular dystrophy
DMT-Cl	4,4'-Dimethoxytrityl chloride
DNG	Deoxyribonucleic guanidine
DNA	2'-deoxyribonucleic acid
dsDNA	double-strand DNA
EDTA	Ethylenediaminetetraacetic acid
EGS	External guide sequence
Et ₃ N	Triethylamine
EtOAc	Ethyl acetate
EtOH	Ethanol
FRET	Fluorescence/Förster Resonance Energy Transfer
FDA	Food and drug administration
G	Guanine
GO	Graphene oxide
gm	Gram
Gn-DNA	2'-Guanidino-DNA
GNA	Glycol nucleic acids
Gu-LNA	2'-Guanidinyl-LNA
Gu-Pet-PNA	Guanidino- <i>N</i> -(pyrrolydiny-2-ethyl) PNA
h	Hours
HBTU	<i>N,N,N',N'</i> -Tetramethyl- <i>O</i> -(1 <i>H</i> -benzotriazol-1-yl)uronium hexafluorophosphate, <i>O</i> -(benzotriazol-1-yl)- <i>N,N,N',N'</i> -tetramethyluronium hexafluorophosphate
HNA	Hexitol nucleic acid
HOBt	Hydroxy benzotriazole
HPLC	High performance liquid chromatography
HRMS	High resolution mass spectrometry
Hz	Hertz
IR	Infrared
L	Liter
LCMS	Liquid Chromatography-Mass Spectrometry
LNA	Locked nucleic acid
LOD	Limit of detection
M	Molar
MALDI-TOF	Matrix assisted laser desorption ionization-time of flight
MBHA	4-(Methyl)benzhydramine
MeOH	Methanol
MF	Molecular formula

mg	Milligram
MHz	Megahertz
min	Minutes
miRNA	MicroRNA
mL	Millilitre
mM	Millimolar
mmol	Millimoles
mmRNA	Mismatch RNA
MS	Mass spectrometry
MOE	Methoxyethyl
MOP	Methoxypropyl
MW	Molecular weight
N	Normal
nm	Nanometer
NMR	Nuclear magnetic resonance
npGNA	N2'-P3' GNA
Obsd.	Observed
ONs	Oligonucleotides
PAGE	Poly acrylamide gel electrophoresis
PCR	Polymerase chain reaction
Pet-ether	Petroleum ether
Pet PNA	<i>N</i> -(pyrrolydiny-2-ethyl) PNA
pi-RNA	Piwi-interacting RNA
pM	Picomolar
PMO	Phosphorodiamidate morpholino oligomers
PNA	Peptide nucleic acid
PNP	<i>para</i> -Nitrophenyl chloroformate
ppm	Parts per million
PS	Phosphorothioates
Py	Pyridine
rC	Ribocytidine
R _f	Retention factor
rG	Riboguanidine
RISC	RNA-induced silencing complex
RNA	Ribonucleic acid
RNAi	RNA interference
R-PCNA	<i>R</i> - Polycarbamate nucleic acid
RP-HPLC	Reverse Phase-HPLC

rt	Room temperature
RT	Reverse Transcriptase
sec	Second
siRNA	Small interfering RNA
SNPs	Single nucleotide polymorphisms
S-PCNA	S- Polycarbamate nucleic acid
SPPS	Solid phase peptide synthesis
ssDNA	Single-stranded DNA
SSOs	Splice-switching oligonucleotides
SVDP	Snake venom phosphodiesterase
T	Thymine
TBA	Thrombin binding aptamer
TBS	<i>tert</i> -Butyldimethylsilyl
TBTU	<i>O</i> -(Benzotriazol-1-yl)- <i>N,N,N',N'</i> -tetramethyluronium tetrafluoroborate
TCA	Trichloroacetic acid
TFA	Trifluoroacetic acid
TFO	Triplex-forming oligonucleotide
TFMSA	Trifluoromethanesulfonic acid
TgO	<i>Thermococcus gorgonarius</i>
THAP	2',4',6'-Trihydroxyacetophenone monohydrate
THF	Tetrahydrofuran
TLC	Thin layer chromatography
T_m	Melting temperature
TMS	Trimethylsilyl
TNA	α -L-Threose nucleic acid
t_R	Retention time
tRNA	Transfer RNA
U	Uracil
UTR	Untranslated region
UV-Vis	Ultraviolet-Visible
w/w	Weight by weight
XRD	X-ray Diffraction

Abstract

The thesis entitled “**Synthesis and Evaluation of Designed (S)-Polycarbamate Nucleic Acids, Erythrulose- based- and 2’- Functionalized- Oligonucleotide Analogs for RNA Targeting**” has been divided into 4 chapters.

Chapter 1: Introduction to nucleic acids

Chapter 2: Applications of synthetic DNA in therapeutics

Section 2A: Design, synthesis and biophysical evaluation of 2’-O-(R/S-serinol) functionalized oligonucleotides

Section 2B: Design and synthesis of 3’/5’-ends serinol capped oligonucleotides and 2’-(ω -O-methylserinol) guarded OMe-RNA/DNA mixmers for efficient splice correction

Section 2C: Design, synthesis and biophysical evaluation of 2’-guanidino functionalized oligonucleotides

Chapter 3: Applications of synthetic DNA mimics in diagnostics

Section 3A: Design and synthesis of (S)-polycarbamate nucleic acids ((S)-PCNA) and their biophysical evaluation

Section 3B: Evaluation of different backbone- and 2’-sugar-modified nucleic acids for DNA diagnostics in conjunction with graphene oxide

Chapter 4: Design and synthesis of erythrulose- based thymine and adenine monomers, and the attempted synthesis of modified oligonucleotides

Chapter 1: Introduction to nucleic acids

This chapter briefly describes structure of nucleic acids, and the applications of the synthetic nucleic acids in antisense therapeutics, graphene oxide based diagnostics and also in understanding the evolutionary chemistry.

The chemical structure of fundamental units of nucleic acids, the different kinds of hydrogen bonding between the canonical nucleobases and the sugar puckering in nucleic

acids are illustrated. The common higher order nucleic acid structures, parameters that control the stability of these structures and their potential in biomedical applications are presented. The principles of the biophysical techniques used to understand the nucleic acids structures (such as UV-vis and circular dichroism spectroscopy) are outlined.

Traditional small molecular drugs target disease proteins such as enzymes and receptors, which represent a small subset of total cellular proteins. In contrast, antisense oligonucleotides (AONs) target pre-mRNA or mRNA- the carriers of genetic information before it is translated into proteins. The RNA therapeutics target mRNA sequences through disruptive or corrective activity with strong and specific interaction of AONs with target mRNA. The principle/mechanism of disruptive antisense approaches (via RNase-H activation, or steric-blocking oligonucleotides, or RNA interference) and corrective antisense approaches (splice correction) are briefly discussed. A plethora of synthetic DNA mimics/analogues are known that are capable of forming strong duplex with the target complementary DNA/RNA and also conferring the necessary metabolic stability. Promising synthetic nucleic acids (like 2'-*O*-sugar modifications, peptide nucleic acids and locked nucleic acids) and their sub-analogues are briefly reviewed.

Despite its short history, graphene oxide (GO), chemically exfoliated from oxidized graphite, has attracted great interest in the area of biomedical application. In view of GO's properties of long-range energy transfer and non-covalent interactions with nucleobases/aromatic compounds via π -stacking, a number of sensors are being developed to detect nucleic acids, metal ions, enzyme activity, and medicines. The principle of nucleic acids sensor based on natural DNA and GO is illustrated with a representative example from the literature.

The emergence of life from the prebiotic environment has remained as an unanswered question since many years. The nucleic acid chemists viewed RNA as the first polymer that emerged as life. The systematic investigation to answer 'why nature has chosen only DNA/RNA as genetic informational carrier/transfer' that sparked the search for the informational polymers that precede RNA is briefly reviewed.

Chapter 2: Applications of synthetic DNA in therapeutics

Section 2A: Design, synthesis and biophysical evaluation of 2'-O-(R/S-serinol) functionalized oligonucleotides

In RNA therapeutics, for the efficient targeting of RNA, the structural design of many sugar-modified AONs involve the 2'-derivatization by electronegative substituents/2'-O-functionalization to retain the N-type sugar conformation, or to lock the sugar in required N-type conformation by a chemical bridge (Figure 1).

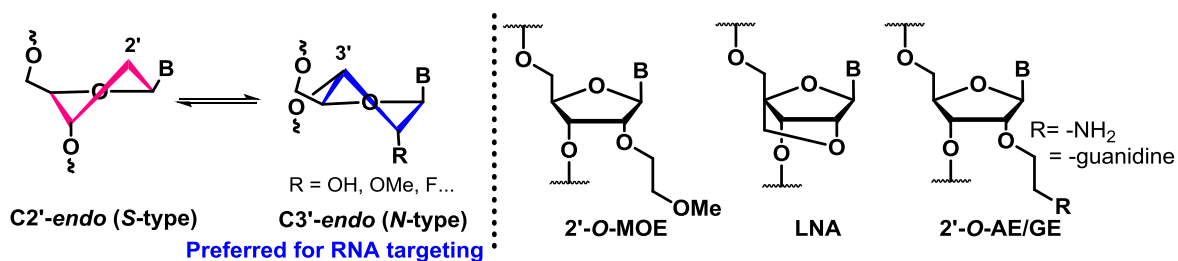


Figure 1 Sugar pucker in nucleic acids and 2'-O-analogues with enhanced affinity to the target and increased enzymatic stability

The 2'-O-aminoethyl/2'-O-aminopropyl derived DNA oligomers form less stable duplexes with complementary RNA, compared to -OMe or MOE-AONs, but show increased stability towards hydrolytic enzymes. Crystallographic study of the 2'-O-aminopropyl- modified AONs bound at the active site of an exonuclease revealed that the protonated amino group of this modification displaced a catalytically important metal cation of the enzyme.

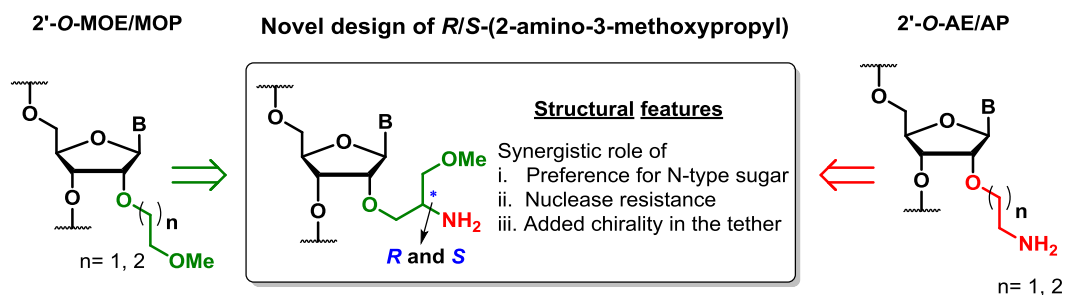
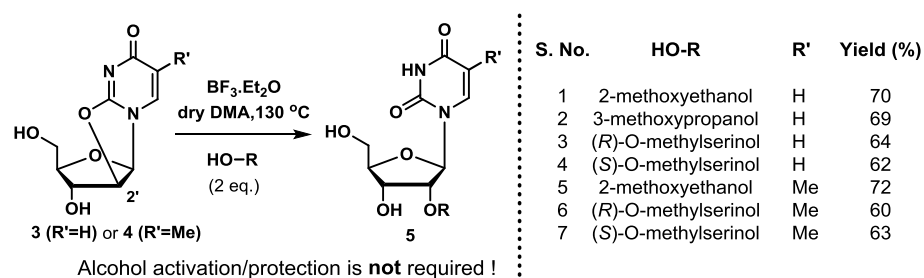


Figure 2 Design of 2'-O-[R/S-2-amino-3-methoxypropyl] modification and the structural features

We designed a novel modification characterized by the presence of a positively charged functionality in conjunction with the methoxy substituent of the 2'-*O*-MOE chain to ensure that the resulting oligomers have advantages of both MOE- and AP-groups (Figure 2). The aminomethoxypropyloxy (AMP) substitution at 2'-position was thus derived from the combination of MOE and AE/AP DNA analogues.

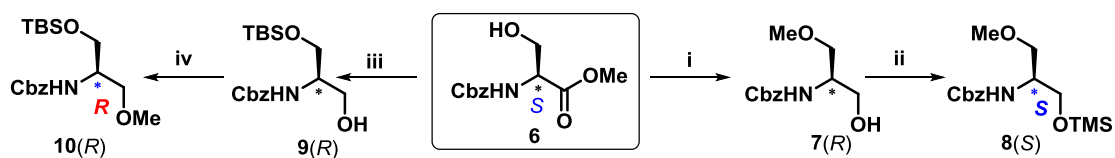
There are previous literature reports to achieve the 2'-*O*-alkylation of uridine using alkyl halides involving extensive protection/deprotection strategies, or by the ring opening of 2,2'-anhydro uridine (**3**) with metal alkoxides, involving large excess of corresponding alcohols, failing which, the alkylation is low yielding. We therefore chose not to adopt them for the present studies.

Scheme 1 Efficient strategy for 2'-*O*-functionalization of pyrimidine-nucleosides



During the synthetic investigations of 2'-*O*-functionalization as reported by Saneyoshi *et al.*, for our designed molecule, we found that even TBS-ethers were equally reactive as TMS-ethers in the ring opening reaction of **3** in presence of $\text{BF}_3 \cdot \text{Et}_2\text{O}$. Consequently, we developed an efficient strategy for 2'-*O*-functionalization of pyrimidine nucleosides with different alcohols through our mono-activation strategy (of electrophile) rather than the reported double-activation strategies and using only two equivalents of the alcohol (Scheme 1).

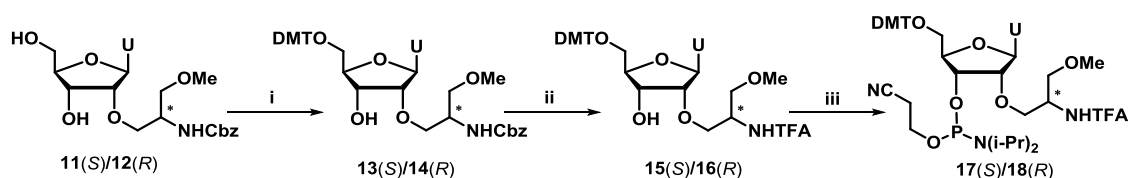
Scheme 2 Synthesis of enantiomeric tethers from *N*-Cbz-L-serine ester **6**



Reagents and conditions: (i) a) Ag₂O, MeI, dry ACN, 12 h; b) NaBH₄, MeOH 0 °C- rt, 83% over two steps (ii) TMS-Cl, Et₃N, dry ACN, 83% (iii) a)TBS-Cl, imidazole, dry DCM; b) NaBH₄, MeOH, 0 °C- rt, 87% over two steps (iv) Ag₂O, MeI, dry ACN, 12 h, 91%.

Stereospecific syntheses of protected [*R/S*-(2-amino-3-methoxy)propanol tethers (*R*-AMP and *S*-AMP) were achieved from L-serine by using a desymmetrization-like concept (Scheme 2). Synthesis of appropriately protected *R*-AMP and *S*-AMP phosphoramidite derivatives of uridine were achieved starting from 2,2'-anhydrouridine **3** (Schemes 1 & 3).

Scheme 3 Synthesis of *R/S*-AMP phosphoramidite derivatives of uridine



Reagents and conditions: (i) DMT-Cl, cat.DMAP, dry pyridine, 89% (ii) a) H₂, Pd/C, 65 psi, MeOH, rt, 6 h; b) Ethyltrifluoroacetate, Et₃N, MeOH, rt, 86-88% over two steps (iii) 2-cyanoethyl-N,N-diisopropylchlorophosphine, DIPEA, dry DCM, 88-92%.

We also synthesized phosphoramidite derivatives of 2'-*O*-MOE (a modification that is an important feature in the antisense drug Mipomersen) and 2'-*O*-MOP uridine derivatives in a cost effective manner, using our improved strategy, as controls in the present study. The phosphoramidites were incorporated into sequence of biological relevance that was used in the splice correction assay, on automated DNA synthesizer to get the modified sequences, Table 1. All the sequences were RP-HPLC purified and MALDI-TOF characterized. The UV-*T_m* studies of the modified oligonucleotides (**ON2-ON9**) revealed that our designed *R/S*-AMP-derived oligonucleotides retained the duplex stability with RNA as observed in 2'-*O*-MOE/2'-*O*-MOP AONs.

Table 1 Modified DNA sequences, their MALDI-TOF mass analyses and biophysical evaluation by UV-*T_m* measurements^a

Sequences/Code ^b	Mass		<i>T_m</i> (°C)	<i>T_m</i> (°C)
	Calcd.	Obsd.	RNA ^c (Δ <i>T_m</i>)	DNA ^d (Δ <i>T_m</i>)
cctcttacctcagttaca ON1	5366.9	5367.2	56.6	54.7
cctcttacctcagtU ^{MOE} aca ON2	5429.9	5430.0	55.8(-0.8)	52.7(-2.0)
cctcttacctcagtU ^{MOP} aca ON3	5440.9	5474.1	55.5(-1.3)	52.7(-2.0)
cctcttacctcagtU ^{S-AMP} aca ON4	5459.6	5460.2	56.2(-0.4)	54.7(0.0)

cctcttacctcagtU ^{R-AMP} aca ON5	5459.6	5460.6	56.2(-0.4)	54.0(-0.7)
cctcttaccU ^{MOE} cagtU ^{MOE} aca ON6	5490.9	5492.7	57.4(+0.8)	49.4(-5.3)
cctcttaccU ^{MOP} cagtU ^{MOP} aca ON7	5515.0	5541.6	56.5(-0.1)	49.4(-5.3)
cctcttaccU ^{S-AMP} cagtU ^{S-AMP} aca ON8	5549.0	5549.8	57.3(+0.7)	53.4(-1.3)
cctcttaccU ^{R-AMP} cagtU ^{R-AMP} aca ON9	5549.0	5551.4	57.5(+0.9)	51.7(-3.0)
tttttttttttttt ON10	3588.1	3589.7		
ttttttttttttU ^{MOE} U ^{MOE} ON11	3708.9	3711.6		
ttttttttttttU ^{MOP} U ^{MOP} ON12	3734.6	3742.3		
ttttttttttttU ^{S-AMP} U ^{S-AMP} ON13	3767.0	3770.1		
ttttttttttttU ^{R-AMP} U ^{R-AMP} ON14	3767.0	3771.3		

a) UV- T_m values were measured by using 1 μ M sequences with 1 μ M cDNA/cRNA in sodium phosphate buffer (0.01M, pH 7.2) containing 100 mM NaCl and are averages of three independent experiments. (Accuracy is ± 0.5 °C). **b)** The lower case letters indicate unmodified DNA, U^{MOE}/U^{MOP} denotes 2'-*O*-methoxyethyl-/2'-*O*-methoxypropyl uridine, U^{S/R-AMP} denotes 2'-*O*-[*S/R*-(2-amino-3-methoxy)propyl] uridine derivatives respectively. **c)** 5'uguaacugagguagagg was the complementary RNA sequence. **d)** 5'tgtaactgaggtaaagagg was the complementary DNA sequence.

To investigate the protection offered by the U^{S-AMP} and U^{R-AMP} tethers against hydrolytic cleavage, we synthesized a homothymine sequence **ON10** and modified it at the 3'-terminus with two consecutive units of U^{MOE}, U^{MOP}, U^{S-AMP} or U^{R-AMP} to get sequences **ON11**, **ON12**, **ON13** and **ON14** respectively (Table 1). These sequences were digested with SVPD under conditions reported earlier. The products of the digestion were analyzed by RP-HPLC and the percent intact AON was plotted against time (Figure 3). The half-life of AONs **ON11** (important modification in the recently marketed Mipomersen) and **ON12** are ~15 min compared to 2 h and 3h of AONs **ON13** and **ON14**, respectively. These results also highlight the introduction of chirality in the minor groove, which did indeed allow differential protection by one of the isomers upon interacting with the active site of the enzyme.

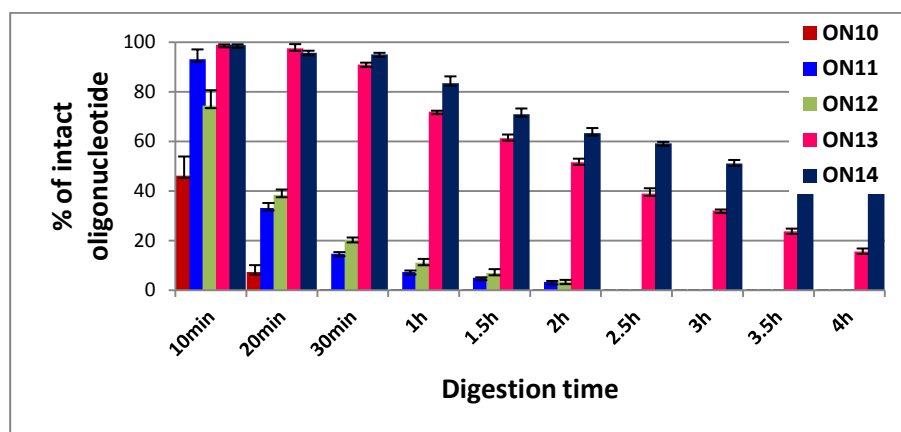


Figure 3 Preferential protection of 3'-end modified AONs

Thus, the 2'-*O*-MOE modification has been transformed into a synergy of chirality, positive charge and hydration module, all in one, in the 2'-*O*-*R/S*-AMP-modified AONs. An efficient synthetic route for 2'-*O*-functionalization of pyrimidine nucleosides through a new mono-activation strategy was accomplished. The syntheses of the two enantiomerically pure tethers were achieved from the same chiral source, L-serine, using a desymmetrization-like concept, *en route* to 2'-*O*-*R/S*-AMP AONs. The demonstrated preserved DNA:RNA duplex stability (like 2'-*O*-MOE, a modification that is being evaluated in clinical trials for several diseases) and the improved stability against hydrolytic enzymes showed that the designed 2'-*O*-*R/S*-AMP AONs have high potential for applications in RNA therapeutics.

Section 2B: Design and synthesis of 3'/5'-ends serinol capped oligonucleotides and 2'-(ω -*O*-methylserinol) guarded OMe-RNA/DNA mixmers for efficient splice correction

The most important factors dictating the potency of the chemical modifications are high binding affinity to the target and high enzymatic stability in biological systems. It has emerged that among the plethora of modified AONs currently under evaluation, the promising AONs have some undesirable drawbacks, e.g., phosphorothioate AONs or OMe/LNA mixmers show non-sequence-specific effects due to non-specific binding to untargeted proteins or due to mismatched non-target recognition due to very high duplex stability. In addition, the non-nucleosidic/nucleosidic chemical modifications are

incorporated into AONs as additional insertions mostly at the 5' or 3' end or both to protect AONs from degradation by exonucleases.

We designed a universal serinol cap for both 3' and 5'-ends to impart enzymatic stability to the AONs. Chemically, our design is a simple three carbon spacer which is spatially equivalent to the three carbons from the sugar unit that are involved in the internucleosidic backbone in DNA/RNA (Figure 4a), but carrying the additional 'amine' functional group which is anticipated to play a crucial role in protection of the AONs from enzymatic degradation (Figure 4b).

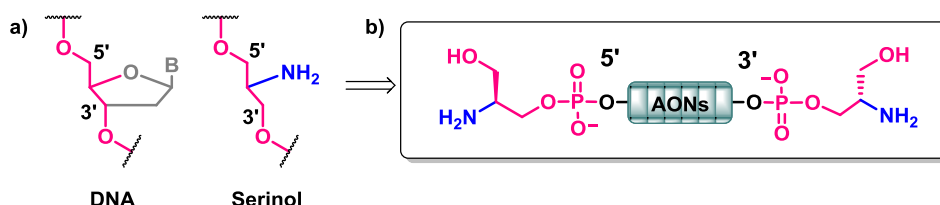
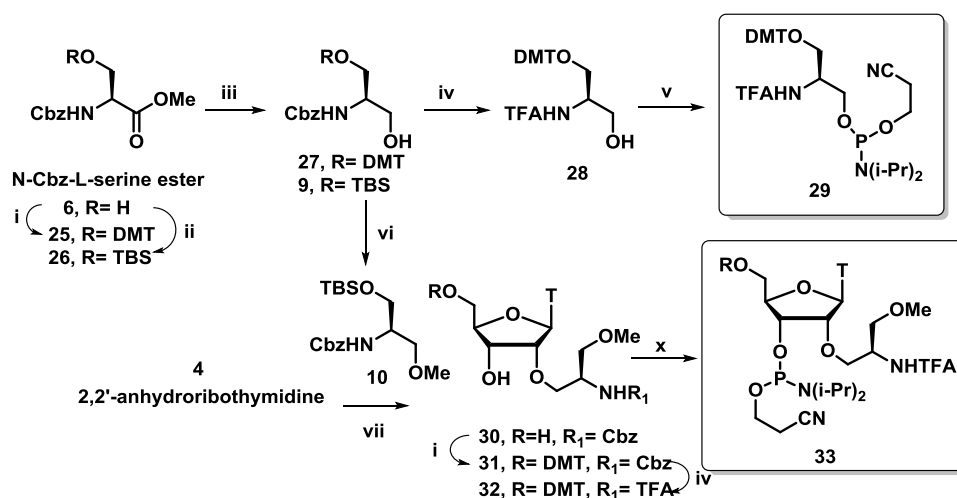


Figure 4 Design of serinol as a universal cap to improve the enzymatic stability of AONs

The synthesis of the universal end-capping monomer **5** and the 2'-O-R-AMP- thymidine monomer (T^{R-AMP}) **10** is outlined in Scheme 4.

Scheme 4 Synthesis of universal capping amidite **5** and T^{R-AMP} amidite **10**



Reagents and conditions: (i) DMT-Cl, cat.DMAP, dry pyridine (ii) TBS-Cl, imidazole, dry DCM (iii) NaBH_4 , MeOH, 0 °C, rt, 78% over two steps for **27** and 87% over two steps for **9** (iv) a) H_2 , Pd/C, 65 psi, MeOH, rt, 4 h; b) CF_3COOEt , Et_3N , MeOH, rt, 86% over two steps (v) 2-cyanoethyl N,N-

diisopropylchlorophosphine, DIPEA, dry DCM, 3 h, 85% (vi) Ag₂O, MeI, dry ACN, rt, 12 h, 91% (vii) **6**, BF₃.Et₂O, dry DMA, 130 °C, 8 h, 65% (viii) 2-cyanoethyl-N,N-diisopropylchlorophosphine, DIPEA, dry DCM, 3 h, 82%.

The phosphoramidites **29** and **33** were incorporated to yield the modified sequences, listed in Table 2. All the sequences were purified by reverse phase HPLC and were characterized by MALDI-TOF mass analysis (Table 2). We studied the protection of these synthesized capped oligomers against enzymatic hydrolytic cleavage *in vitro* and compared the results with unmodified **ON1** and **ON17**. After digestion with SVPD for 24h, subsequent analysis of the digestion products by RP-HPLC revealed that both the capped DNA sequences **ON15** and **ON16** were found to resist hydrolysis to some extent and degraded by 70% and 50%, respectively, while the native **ON1** was completely digested after ~30 min. (Figure 5a). The protective effect was more pronounced in 2'-OMe RNA, where the terminal serinol units were found to be able to completely guard the 2'-OMe-RNA oligomers (**ON18** and **ON19**) against hydrolytic cleavage by SVPD for up to 24 h compared with the uncapped OMe-RNA sequence (<10% intact after 5h) (Figure 5a).

Table 2: The sequences synthesized, their MALDI-TOF mass analyses and UV-*T_m* values with complementary RNA and RNA with single mismatch (top panel); commercially purchased sequences used in the study (bottom panel)

Sequence ^a (Code)	Mass Calcd/Obsd	UV- <i>T_m</i> °C ^b cRNA/mmRNA
CCTCTTACCTCAGTTACA (ON1)	5366.9/5367.2	56.6/50.4
SCCTCTTACCTCAGTTACAS (ON15)	5672.9/5690.3	58.3/50.4
SSCCTCTTACCTCAGTTACASS (ON16)	5979.0/6002.0	58.4/48.7
ccucuuaccucaguuaca (ON17)	5823.0/5824.4	77.0/66.2
SccucuuaccucaguuacaS (ON18)	6129.0/6151.4	78.2/67.9
SSccucuuaccucaguuacaSS (ON19)	6435.1/6453.0	76.8/65.4
SSccxcu _x accxcaguxacaSS (ON20)	6786.6/6787.8	76.4/64.0
cC ^L ucT ^L uA ^L cC ^L ucA ^L guT ^L aC ^L a (ON21)		
cT ^L cuC ^L uC ^L aC ^L caT ^L ugA ^L cA ^L a (ON22)		

^aa/c/g/u are RNA or 2'-OMeRNA nucleotides; S is a serinol capping unit; x= T^{R-AMP} is a 2'-O-(2-amino-3-methoxypropyl)ribothymidine unit; C^L, A^L and T^L are the conventional locked nucleotides

^bUV-*T_m* values were measured by using 1μM sequences with 1μM cDNA/cRNA in sodium phosphate buffer (0.01M, pH 7.2) containing 100 mM NaCl and are averages of three independent experiments. (Accuracy is ±0.5 °C).

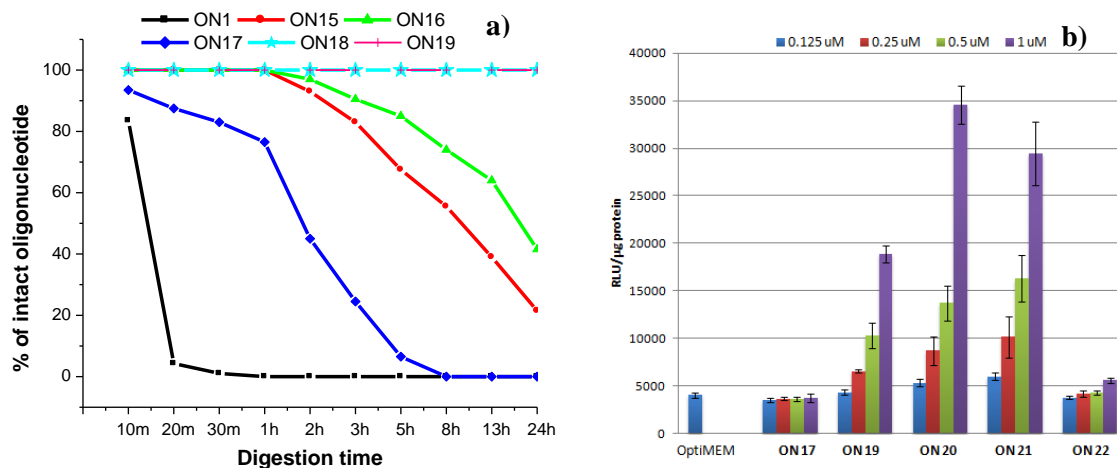


Figure 5 SVPD digestion of 5'/3'-end capped oligonucleotides compared to native DNA (a) and 2'-OMe and the concentration dependent splice correction activity of oligonucleotides (b).

Cell lines expressing luciferase pre-mRNA, interrupted by an aberrantly spliced β -globin intron, HeLa pLuc705, were used to monitor the splice-switching activity of the modified oligonucleotides. The results shown in Figure 5b clearly indicate that the dose-dependent luciferase expression in the cells treated with serinol-end-capped **ON19** was better than that for the uncapped OMe-RNA **ON17**. The expression was found to be even better for the sequence **ON20**, which comprised four T^{R-AMP} units (i.e. $\sim 20\%$ substitution of U^{OMe} with $2'-O-T^{R-AMP}$), in addition to serinol end-caps. The expression of corrected fluorescent protein was found to be as good as or even slightly better than with an LNA/OMe mixmer containing about $\sim 40\%$ LNA substitution. We have also shown by UV- T_m studies, that the single mismatch discrimination is as high as that shown by OMe-RNA oligomer while forming a duplex with the target RNA (Table 2) and therefore off-target effects may also be expected to be less compared with the LNA-OMe mixmer sequences, which sometimes tolerate mismatched bases in the target RNA.

In summary, we described the simple design and synthesis for an efficient protection of AONs. To harness the maximum benefit of our serinol-derived modifications, we incorporated our nucleosidic and non-nucleosidic serinol derivatives in 2'-OMe RNAs without disturbing the efficacy or fidelity of duplex formation. The 5'- and 3'-end capping of oligonucleotides with serinol provided excellent protection against SVPD even after 24 h

digestion (30%, 50%, 100% and 100% protection for AONs **ON15**, **ON16**, **ON18** and **ON19** respectively). We showed that the 3'- and 5'- capped 2'-OMe-AON with ~20% evenly dispersed modified T^{R-AMP} units (**ON20**) is as effective as an LNA-OMe mixmer containing ~40% LNA (**ON21**) in steric blocking splice correction of an aberrant β -globin gene.

Section 2C: Design, synthesis and biophysical evaluation of 2'-guanidino functionalized oligonucleotides

For therapeutic applications to target any DNA/RNA with enhanced affinity, an obvious/easy option is to reduce the inter-strand repulsions, achieved through replacement of charged phosphate backbone by a neutral backbone (like peptide, carbamate, etc.,) or by introducing functional groups that are cationic in nature at physiological pH (amine/guanidine through phosphate backbone, 2'-O-tether or as a tether in nucleobase). These zwitter-ionic modified oligonucleotides showed properties of enhanced binding affinity to complementary strands, high nuclease resistance and enhanced cellular uptake.

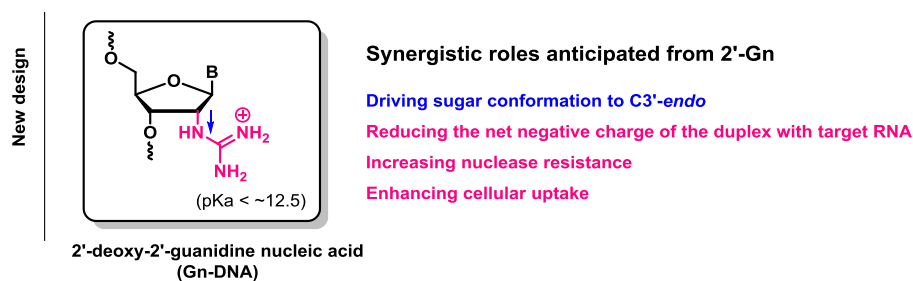


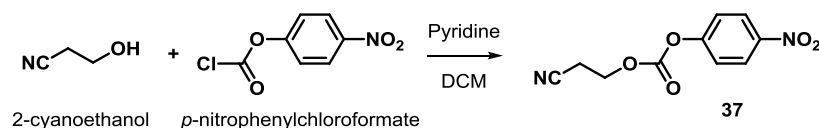
Figure 6 Our design of Gn-DNA towards potential antisense applications

We designed to substitute the 2'-position of the deoxyribose sugar of DNA with the guanidine group (Figure 6). Once protonated at physiological pH, the guanidinium group can act as an electron withdrawing group and hence may affect the ring puckering of ribose sugar, like 2'-OMe or 2'-F substituted oligonucleotides, towards C3'-endo conformation required for an efficient targeting of RNA.

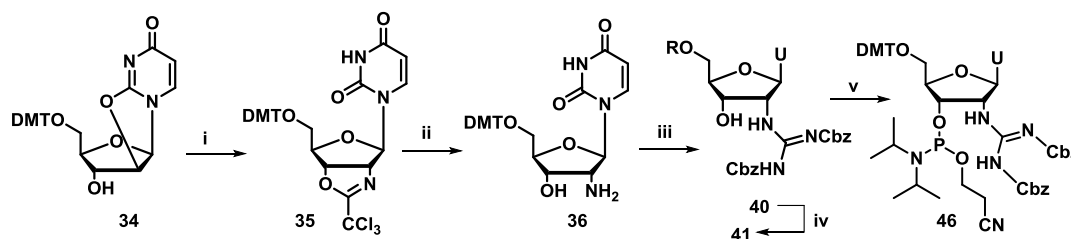
Towards the synthesis of 2'-Gn nucleoside, we envisaged guanidinylation of 2'-deoxy-2'-amino-ribonucleoside using commercially available reagents. We synthesized a new reagent, 2-cyanoethyl(4-nitrophenyl)carbonate **37** for -CEOC protection of the

guanidinylation agent using relatively cheap *p*-nitrophenylchloroformate rather than disuccinimidyl carbonate. However, the 2'-guanidinylation of compound **36** was unsuccessful using mono-CEOC-protected guanidinylation reagents. Successful guanidinylation of **36** to furnish the compound **40** was achieved using *N,N'*-bis(benzyloxycarbonyl)-1*H*-pyrazole-1-carboxamide, scheme 5.

Scheme 5 Synthesis of 2-cyanoethyl(4-nitrophenyl)carbonate reagent **37** for -CEOC protection of guanidinylation reagents



Scheme 6 Synthesis of 2'-guanidine-derived phosphoramidite derivative **46**



Reagents and conditions: (i) CCl_3CN , Et_3N , reflux (ii) 6*N* NaOH/EtOH , 16 h, 68% yield over two steps (iii) *N,N'*-bis(benzyloxycarbonyl)-1*H*-pyrazole-1-carboxamide, DIPEA, dry DMF, 4 h, 71% (iv) 10% Pd/C, H_2 (50 psi), MeOH, 2 h (v) 2-cyanoethyl-*N,N*-diisopropylchlorophosphine, DIPEA, dry DCM, 3 h, 86%.

The ^1H NMR analysis revealed a significant conformational change when 2'-amino group (based on, %S = $100 \times (J_{\text{H1}'\text{-H2}' - 1})/6.9$) was converted to the 2'-Gn group, Table 3.

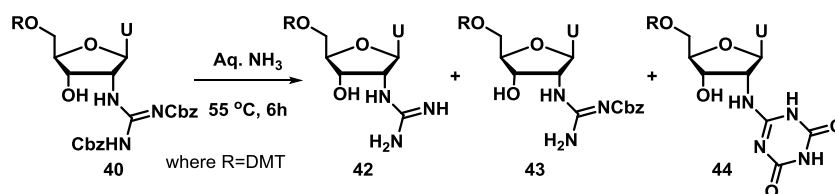
Table 3 %S & %N calculations based on $J_{\text{H1}'\text{-H2}'}$ of 2'-amine (**36**) and 2'-guanidine (**41**) derivatives

Compound	$J_{\text{H1}'\text{-H2}'}$ (Hz)	%S (C2'-endo)	%N (C3'-endo)
36 -NH ₂	6.65	82	18
41 -NH-C(=NH)-NH ₂	2.25	18	82

We examined the stability of di-Cbz-guanidine derivative **40** to the final cleavage conditions of oligonucleotides from CPG-resin (aq. NH_3 , 55 °C, 6 h). HRMS analysis of the

reaction mixture after aqueous ammonia treatment of monomer **40** for 6 hours at 55 °C, showed the exclusive formation of the mono-Cbz protected guanidine **43** along with traces of free guanidine **42** and 2'-N-linked triazine derivative **44**.

Scheme 7 Stability of **40** to final ammonia-deprotection conditions



The amidite **46** was incorporated by solid phase synthesis into the DNA sequences containing single and double modification (Table 4, **ON23-ON26**). The modified oligonucleotides were cleaved from the resin and were purified by HPLC. Surprisingly, MALDI-ToF analysis of the products obtained by HPLC purification, confirmed the formation of modified oligonucleotides **ON23/ON25** and **ON24-ON26** containing 2'-N-Cbz-Gn (**X**, corresponding to **43**) and, 2'-N-triazine (**Y**, corresponding to **44**) respectively, with no traces of 2'-Gn-containing oligonucleotides. The UV-*T_m* studies of the modified nucleotides revealed the destabilization of the duplexes with cDNA and cRNA and the destabilization is less pronounced in 2'-N-Cbz-Gn-DNAs.

Table 4 2'-Gn derived oligonucleotides, MALDI-ToF characterization and their UV-*T_m* data.

Code	Sequences (5'-3') ^a	Mass Calcd/Obsd	UV <i>T_m</i> (ΔT_m) in °C ^b	
			cDNA	cRNA
ON1	CCTCTTACCTCAGTTACA	5366.9/5367.2	54.7	56.6
ON23	CCTCTTACCTCAGT <u>X</u> ACA	5543.9/5543.0	51.1(-3.6)	54.1(-2.5)
ON24	CCTCTTACCTCAGT <u>Y</u> ACA	5478.9/5474.6	50.3(-4.4)	53.0(-3.6)
ON25	CCTCTTACC <u>X</u> CAGT <u>Y</u> ACA	5655.9/5654.0	45.6(-9.1)	49.4(-7.2)
ON26	CCTCTTACC <u>Y</u> CAGT <u>Y</u> ACA	5590.9/5585.5	44.4(-10.3)	48.4(-8.2)
ON17	ccucuuaccucaguuaca	5823.0/5824.4	53.4	77.0
ON27	ccucuuaccucagu <u>X</u> aca	5984.0/5980.4	49.7(-3.7)	72.7(-4.3)
ON28	ccucuuaccucagu <u>Y</u> aca	5919.0/5932.0	45.2(-8.2)	70.7(-6.3)

^aX=2'-(*N*-Cbz)-Gn and Y=2'-*N*-triazine; ^bUV- T_m values were measured by annealing 1 μ M each sequence with 1 μ M cDNA in sodium phosphate buffer (0.01M, pH 7.2) containing 100mM NaCl and are averages of three independent experiments. Lower case represents 2'-OMe nucleotides

In summary, the guanidinylation of 2'-deoxy-2'-aminouridine, though proved to be challenging, was achieved with *N,N'*-bis(benzyloxycarbonyl)-1*H*-pyrazole-1-carboxamide in moderate yield. Our design of 2'-Gn modification to play a synergistic role resulted in the serendipitous formation of the 2'-*N*-triazine derivative which has potential as a double-headed nucleoside. The 2'-*N*-triazine and 2'-(*N*-Cbz)-Gn modification-containing oligonucleotides destabilize the duplexes with cDNA and cRNA, the degree of destabilization being less in (*N*-Cbz)-Gn oligonucleotides, probably due to the stabilizing stacking interactions offered by the -Cbz group.

Chapter 3: Applications of synthetic DNA mimics in diagnostics

Section 3A: Design and synthesis of (*S*)-polycarbamate nucleic acids ((*S*)-PCNA) and their biophysical evaluation

To address the limitations of PNA such as poor aq. solubility, cellular uptake and ambiguity in mode of binding, our group recently reported the simple variation in the structure of PNA leading to *R*-polycarbamate nucleic acids (*R*-PCNA). The linking of individual enantio-pure monomers derived from L-serine, *via* carbamate linkage to form *R*-PCNA (Figure 7) conserved six atom repeating backbones of DNA/PNA. The nucleobase attachment to the carbamate backbone *via* a secondary amide function could circumvent the likelihood of *cis/trans* isomers present in tertiary amide linker groups in PNA.

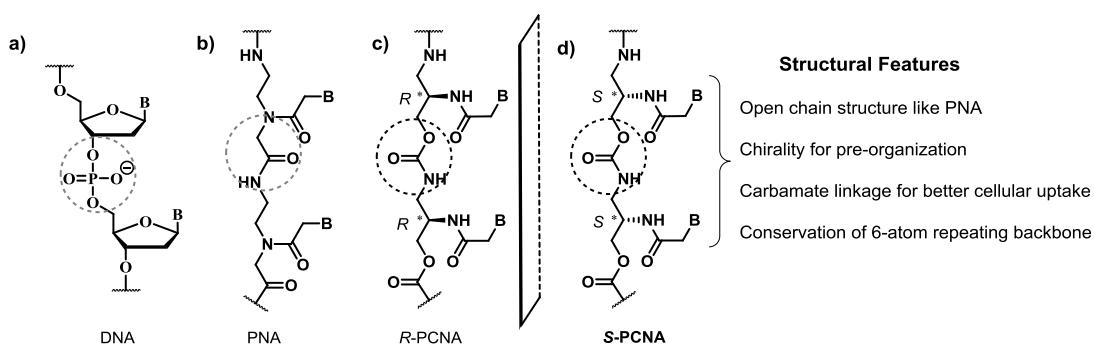
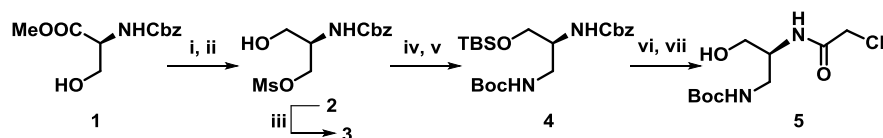


Figure 7 Chemical structures of DNA, PNA and our designed polycarbamate nucleic acids

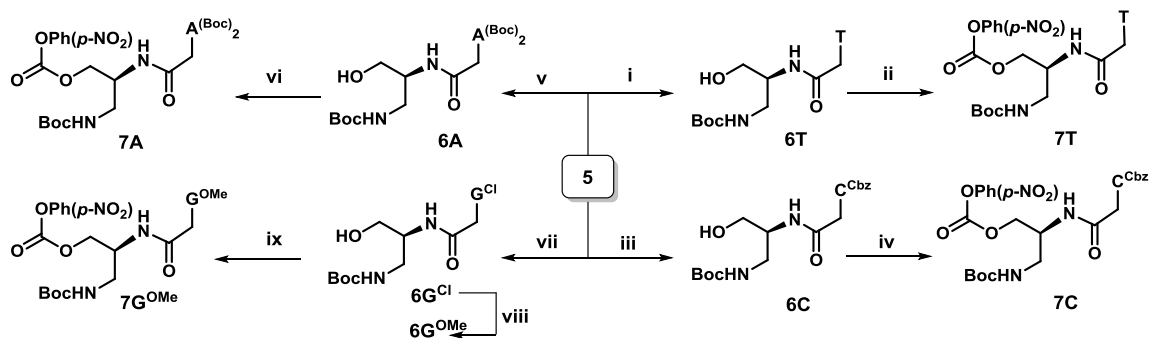
Synthesis of *S*-PCNA monomers (T, C, G and A) and *R*-PCNA monomer (T) for Boc-strategy starting from L-serine were accomplished in enantiomerically pure form, outlined in Schemes 8, 9, and 10.

Scheme 8 Synthesis of *S*-PCNA precursors for Boc-strategy



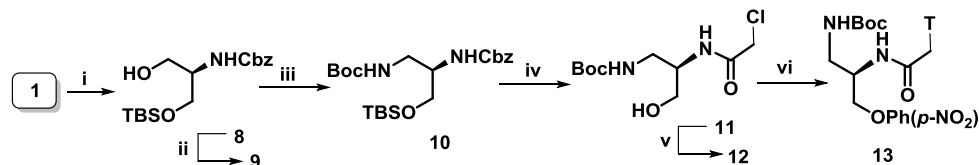
Reagents and conditions: (i) Ms-Cl, Pyridine, 0 °C, dry DCM, 15 min (ii) NaBH₄, MeOH, 0 °C-rt, 75% over two steps (iii) TBS-Cl, imidazole, dry DCM, 5 h, 95% (iv) NaN₃, dry DMF, 60 °C, 76% (v) Raney Ni, H₂, 40 psi, (Boc)₂O, EtOAc, 4 h, 84% (vi) H₂, Pd/C MeOH, 50 psi, 6 h (vii) ClCOCH₂Cl, K₂CO₃, Dioxane:H₂O (1:1), 8 h, 80% over two steps.

Scheme 9 Synthesis of *S*-PCNA carbonate monomers carrying genetic alphabets T, C, A and G



Reagents and conditions: (i) Thymine, K₂CO₃, dry DMF, rt, overnight, 91% (ii) *p*-NO₂-PhOCOCl, dry DCM, pyridine, 70% (iii) Cbz-cytosine, K₂CO₃, dry DMF, overnight, 80% (iv) *p*-NO₂-PhOCOCl, dry DCM, pyridine, 72% (v) (Boc)₂-Adenine, K₂CO₃, dry DMF, overnight, 75% (vi) *p*-NO₂-PhOCOCl, dry DCM, pyridine, 86% (vii) 2-amino-6-chloropurine, K₂CO₃, dry DMF, overnight, 80% (viii) 0.7 M NaOH, MeOH, rt, 10 h, 85% (ix) *p*-NO₂-PhOCOCl, dry DCM, pyridine, 66%.

Scheme 10 Synthesis of T-monomer for *R*-PCNA synthesis via Boc-strategy



Reagents and conditions: (i) a) TBS-Cl, imidazole, dry DCM, 5 h, 95% b) NaBH₄, MeOH, 0 °C-rt, 81% over two steps (ii) Ms-Cl, Pyridine, 0 °C, dry DCM, 15 min (iii) a) NaN₃, dry DMF, 60 °C, 76% b) Raney Ni, H₂, 40 psi, (Boc)₂O, EtOAc, 4 h, 87% (iv) a) H₂, Pd/C MeOH, 50 psi, 6 h b) ClCOCH₂Cl, K₂CO₃, Dioxane:H₂O (1:1), 8 h, 80% over two steps (v) Thymine, K₂CO₃, dry DMF, rt, overnight, 88% (vi) *p*-NO₂-PhOCOCl, dry DCM, pyridine, 72%.

Polycarbamate nucleic acids were synthesized on solid support and cleaved from the support under strong acidic conditions, RP-HPLC purified and their melting studies of the duplexes with cDNA were performed, Table 5.

Table 5 UV-melting studies of *S*-PCNA oligomers

Code	<i>S</i> -PCNA oligomers	UV- T_m °C* cDNA
<i>S</i>-PCNA1	H'-TTTTT TTT-LysNH ₂	56.2
<i>R</i>-PCNA1	H'-TTTTT TTT-LysNH ₂	50.1
<i>S</i>-PCNA2	H'-CTTCTTCCTT-LysNH ₂	57.6
<i>S</i>-PCNA3	H'-CATTGTCACACT-LysNH ₂	58.9 (12mer:18mer)
DNA4	5' TTTTTTTT 3'	21.5
DNA5	5' CTTCTTCCTT 3'	31.2
DNA6	5' CACCATTGTCACACTCCA 3'	53.9 (18mer:18mer)

*UV- T_m values were measured by annealing 1 μ M sequences with 1 μ M cDNA in sodium phosphate buffer (0.01M, pH 7.2) containing 10mM NaCl, 0.1mM EDTA and is an average of three independent experiments.

Lane **I**: ***R*-PCNA1:DNA1** complex

Lane **II**: ***S*-PCNA1:DNA1** complex

Lane **III**: Single stranded **DNA1**

Lane **IV**: **DNA4:DNA1** complex



Figure 8 Ethidium bromides staining of 15% non-denaturing PAGE analysis of complexes of ***S*-PCNA1**, ***R*-PCNA1** and **DNA4** with complementary **DNA 1**

We have successfully synthesized the enantiomeric *S*-PCNA oligomers from the same chiral source L-serine, as that of *R*-PCNA, by simply using the same set of chemical reactions, the only change being in the order of the chemical reactions. We performed the UV- T_m studies and observed that these *S*-PCNA oligomers also bind to DNA, like *R*-PCNA oligomers, supported by polyacrylamide gel electrophoresis. Binding with RNA was not observed.

Section 3B: Evaluation of different backbone- and 2'-sugar-modified nucleic acids for DNA diagnostics in conjunction with graphene oxide

The basic concept of all the Graphene Oxide (GO)-based probes takes advantage of GO's preferential interaction with ssDNA over dsDNA. The exposed nucleobases in ssDNA are adsorbed strongly on the GO surface through strong π -stacking in contrast to the nucleobases in dsDNA. The stacked pairs of nucleobases in dsDNA are effectively hidden in the helical structure, which prevents the direct interaction of nucleobases with GO surface. Therefore, when the fluorescently labeled ssDNA probes bound to the GO surface hybridize with their complementary target DNA, forming a duplex, it is desorbed from the GO surface, leading to a recovery of fluorescence (Figure 9, Path A). The phosphodiester linkages in the probe as well as target allow the reversible adsorption/desorption process and would lead to unwanted Path B shown in Figure 9.

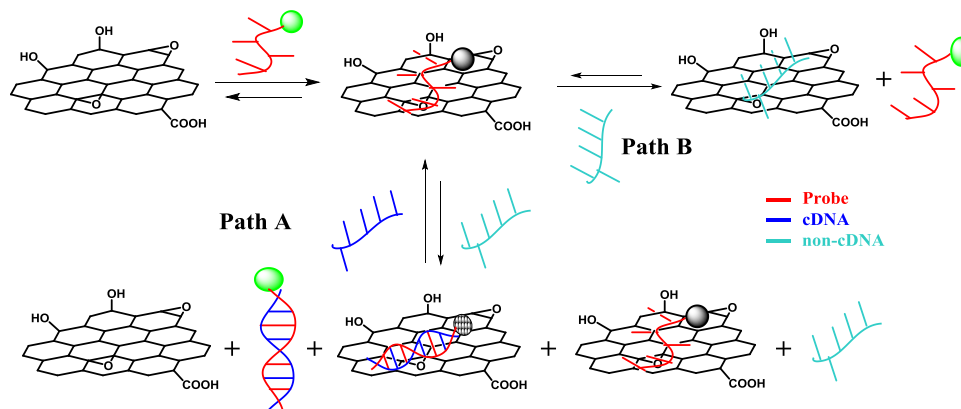


Figure 9 Schematic representation of fluorescence quenching of labeled probes with GO and the possible pathways in restoring the fluorescence by cDNA/random DNA.

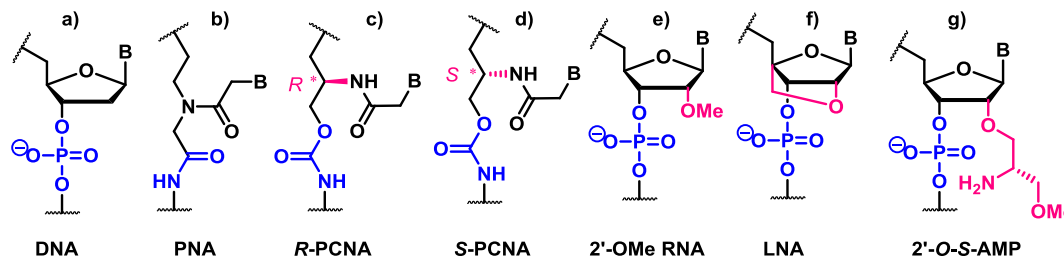


Figure 10 DNA and selected chemically modified nucleic acids that exhibit strong duplexation abilities and/or are expected to have additional electrostatic interactions with the probe.

This section deals with the understanding of interactions of different backboned octathymine probes [namely phosphate- (DNA), polyamide- (PNA) and polycarbamate- (PCNA), Figure 10a-d] with GO using fluorescence quenching and CD studies and also the fluorescence restoration with cDNA. The similar studies were extended to mixed pyrimidine/purine sequences of 10-mer and 12-mer length. The thermal desorption studies of the probes from GO surface were also performed. Further we studied the effect of sequence length on the desorption of modified DNA duplexes (Figure 10e-g).

The fluorescence of neutral backboned labeled probes [with Carboxyfluorescein (CF)], PNA-CFs and *R/S*-PCNA-CFs oligomers was quenched instantly whereas the fluorescence quenching of DNA-CFs was found to be much less. There was almost no or very little fluorescence increase observed upon addition of cDNA to (**I-CF** to **III-CF**):GO) and (**VII-CF** to **XI-CF**):GO, whereas **IV-CF**:GO showed specific and measurable (40%) restoration of the fluorescence intensity upon addition of cDNA **V** (Figure 11). In case of DNA-CFs the cDNA-induced restoration was proved to be non-specific, as similar restoration was observed even with random DNA.

Table 6 Sequences synthesized, MALDI-TOF characterization and UV- T_m data

Description (length)	Sequence-Code (5'/N-3'/C)	Mass Calcd./Obsd.	T_m °C* cDNA
DNA (8)	ttttttt I-CF	2908.6/2907.2	~10
PNA (8)	ttttttt II-CF	2575.9/2579.6	45.0
<i>R</i> -PCNA (8)	ttttttt III-CF	2759.9/2763.0	50.1
<i>S</i> -PCNA (8)	ttttttt IV-CF	2759.9/2761.6	56.2
cDNA	gcaaaaaaacg V	3680.5/3676.7	-
non-cDNA	agtgacctac VI	3012.0/3018.4	-
DNA (10)	cttctcctt VII-CF	3457.0/3459.2	29.1
PNA (10)	cttctcctt VIII-CF	3106.4/3108.6	52.0
<i>S</i> -PCNA (10)	cttctcctt IX-CF	3267.9/3268.9	57.6
<i>S</i> -PCNA (12)	atattattaatt X-CF	3936.2/3943.2	56.8
<i>S</i> -PCNA (12)	actaaaactcat XI-CF	3900.5/3907.5	59.2
cDNA	aaggaagaag XII	3134.1/3140.1	-
cDNA	aattaataatat XIII	3651.5/3673.2	-

cDNA	atgagtttagt XIV	3690.5/3705.5	-
DNA (18)	cctcttacctcagttaca XV-CF	5904.0/5904.4	54.8
2'-OMe RNA (18)	ccucuaccucaguuaca XVI-CF	6360.1/6360.9	53.4
DNA, $\underline{x} = \text{t}^{\text{LNA}}$ (18)	cc \underline{x} ct \underline{x} acc \underline{x} cagt \underline{x} aca XVII-CF	6016.0/6015.8	64.1
DNA, $\underline{x} = \text{u}^{\text{S-AMP}}$ (18)	cc \underline{x} ct \underline{x} acc \underline{x} cagt \underline{x} aca XVIII-CF	6260.2/6261.8	46.1
cDNA	tgtaactgaggtgaagagg XIX	5624.9/5623.9	-

Thermal desorption studies revealed that the synthesized probes **I-CF** to **XI-CF** desorbed from the GO surface approximately at the same temperature independent of the presence of cDNA, except **IV-CF**. Only **IV-CF** has significant difference in desorption temperature (-8.2 °C) in the presence and absence of cDNA (38.4 °C and 46.6 °C respectively). Our experimental results of cDNA-induced desorption with DNA and probes with different backbones such as PNA, *R*-PCNA and *S*-PCNA therefore may be explained based on these results. The fluorescence restoration of **XV-CF** to **XVIII-CF** from GO surface upon challenging with cDNA **XIX** was not prominent; the trend was dominated by the **XVII-CF** probe, which could be easily attributed to high duplex stability imparted by the LNA-modifications in this probe. The dominance effect by **XVII-CF** was more pronounced when the system is incubated at 50 °C for 30 min (Figure 12).

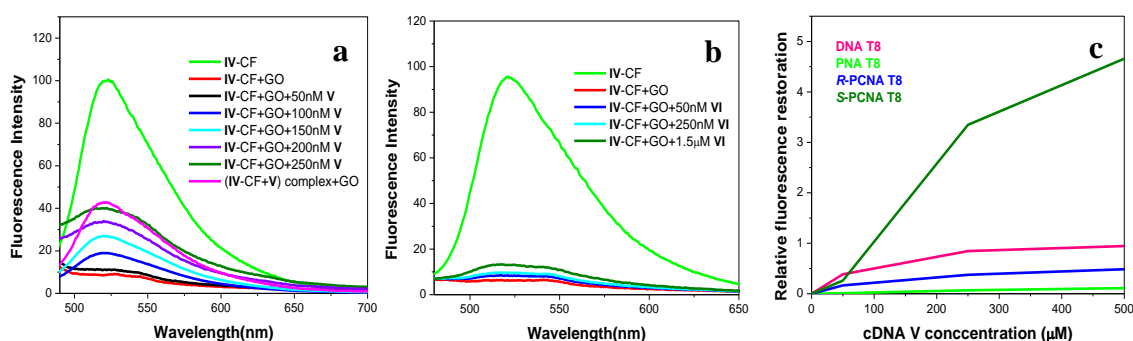


Figure 11 Significant, specific and relative fluorescence restoration of **IV-CF** with cDNA **V**

In summary, for the first time we have introduced the use of synthetic analogues in conjunction with GO to develop nucleic acid-based capture probes. cDNA-induced specific desorption of the probes from the GO surface was performed. Among the varied backbone T₈-CF sequences, only the *S*-PCNA-T₈-CF (**IV-CF**) showed the selective prominent and

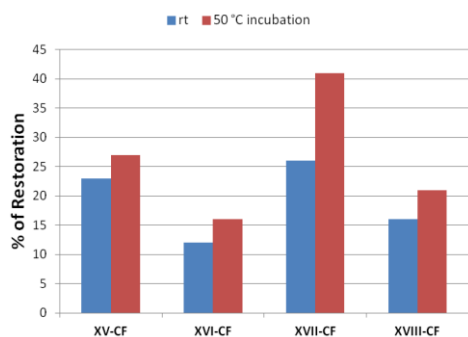


Figure 12 cDNA induced relative % fluorescence restoration at room temperature and 50 °C incubation for 30 minutes

specific fluorescence restoration upon the cDNA (V) addition. Thermal desorption of the synthesized short sequence probes from the GO surface were studied and are in good accordance with the cDNA-induced spontaneous desorption at room temperature. Additionally, we synthesized new DNA probes containing modifications of *S*-AMP design and LNA. However the fluorescence restoration studies need to be optimized. In conclusion, our studies

illustrate the possibility of using novel, stable synthetic DNA analogues and will help in better understanding of the nucleic acids interaction with GO, stimulating further research in this important area.

Chapter 4: Design and synthesis of erythrose- based thymine and adenine monomers, and the attempted synthesis of modified oligonucleotides

Since the discovery of the double helical DNA structure, Watson-Crick base pairing has been widely speculated to be the likely mode of both information conservation and communication in the earliest genetic polymers. In spite of the discovery of many other RNA enzymes with functions relevant to the RNA world, it is still difficult to imagine RNA as the first genetic polymer, because it is difficult for an inherently unstable molecule as complex as RNA to self-assemble from the potential prebiotic molecules and to remain long enough in that environment to carry out the initial processes of information transfer and catalysis. These difficulties led to another hypothetical period referred as ‘Pre-RNA world’. Among the plethora of synthetic nucleic acids available, very few have been found with efficient template assisted polymerization such as glycol nucleic acids (GNA), threose nucleic acids (TNA), hexitol nucleic acids (HNA) and peptide nucleic acids (PNA) as potential pro-genitors of RNA (Figure, 13)

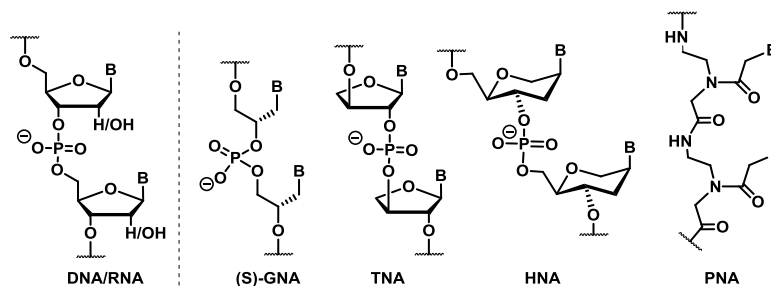


Figure 13 Chemical structures of natural genetic polymers (DNA/RNA) and recognized Pre-RNA synthetic genetic polymers

We designed a new acyclic modification based on the four carbon carbohydrate L-erythrulose, a ketose, perhaps the first of its kind in the modified nucleic acids, in contrast to the aldose in Eschemoser's TNA. In our design, the nucleobases are attached to the backbone through a methylene linker to the sp^2 carbon (carbonyl group of ketose) and individual units are linked through vicinal phosphodiester linkages between 3' and 4' of successive units (Figure 14). The backbone is shortened by one-atom relative to DNA/RNA and is comparable to that in TNA and GNA. Our design possesses more rigidity compared to the GNA and more flexibility compared to the TNA (a cyclic modification). It is of interest to note that both TNA and S-GNA have been proven to form stable hetero-duplexes with RNA.

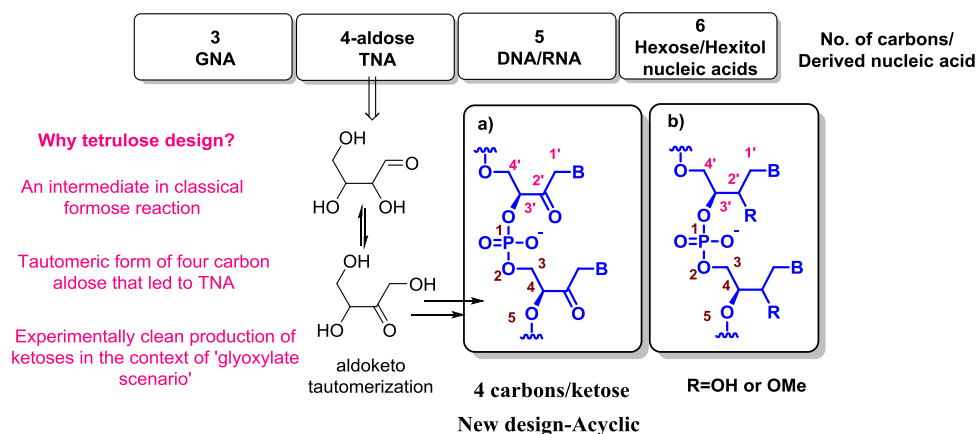
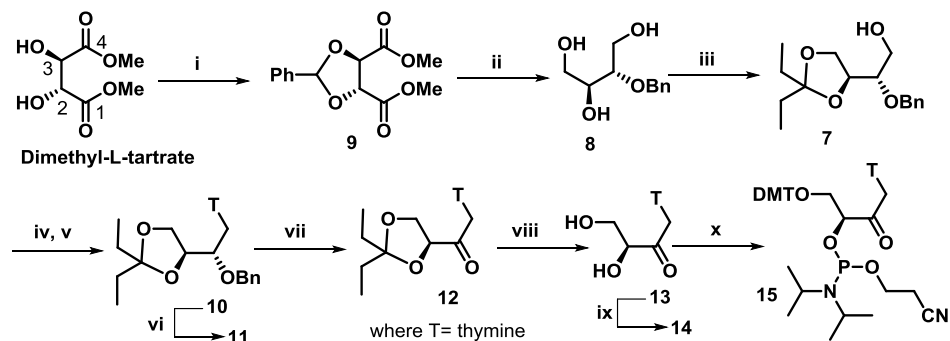


Figure 14 Design of 4-carbon ketose, erythrulose-based nucleic acids

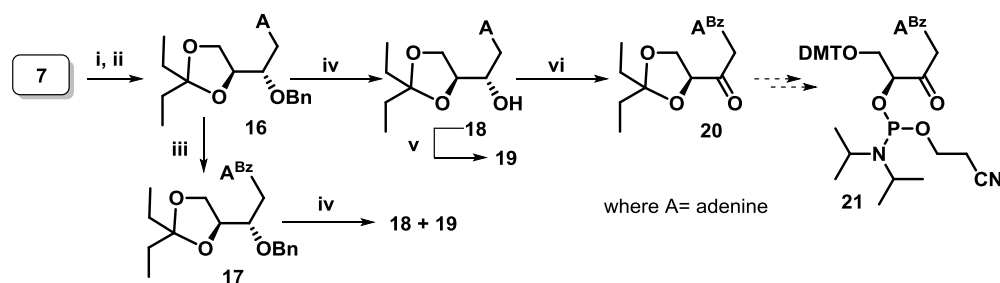
Synthesis of desired thymine and adenine monomers were accomplished starting from dimethyl-L-tartrate (Scheme) rather from L-erythrulose (low yielding).

Scheme 11 Synthesis of erythrulose-based thymine monomer from dimethyl-L-tartrate



Reagents and conditions (i) $\text{PhCH}(\text{OMe})_2$, under reduced pressure (ii) LiAlH_4 , AlCl_3 , Et_2O , dry DCM (iii) 3-pentanone, dry THF, p-TSA (iv) MsCl , Et_3N , dry DCM (v) NaH , dry DMF, thymine (vi) 10% Pd/C , ammonium formate, MeOH , reflux (vii) CrO_3 , Ac_2O , Pyridine, dry DCM (viii) 5% aq. H_2SO_4 , MeOH (ix) DMT-Cl , Pyridine, cat. DMAP (x) 2-cyanoethyl-N,N-diisopropylchlorophosphine, DIPEA , dry DCM.

Scheme 12 Synthesis of erythrulose-based adenine monomer from dimethyl-L-tartrate



Reagents and conditions (i) MsCl , Et_3N , dry DCM (ii) NaH , dry DMF, Adenine (iii) a) Bz-Cl , pyridine, rt, overnight; b) NH_4OH , MeOH , rt, 30 min (iv) 10% Pd/C , ammonium formate, MeOH , reflux, 1 h (v) a) TMS-Cl , Pyridine; b) Bz-Cl , NH_4OH (vi) CrO_3 , Ac_2O , Pyridine, dry DCM.

In order to examine the stability offered by the new erythrulose-based thymine monomer, we attempted to synthesize the modified oligonucleotides, Table 7. Unfortunately the MALDI-TOF analysis of the modified oligonucleotides revealed mass difference of ~220-290 Daltons. The mass values could not be attributed by us to any possible modified oligonucleotides resulting from unwanted reactions (probably at the carbonyl) during the sequence of chemical reactions on automated DNA synthesizer.

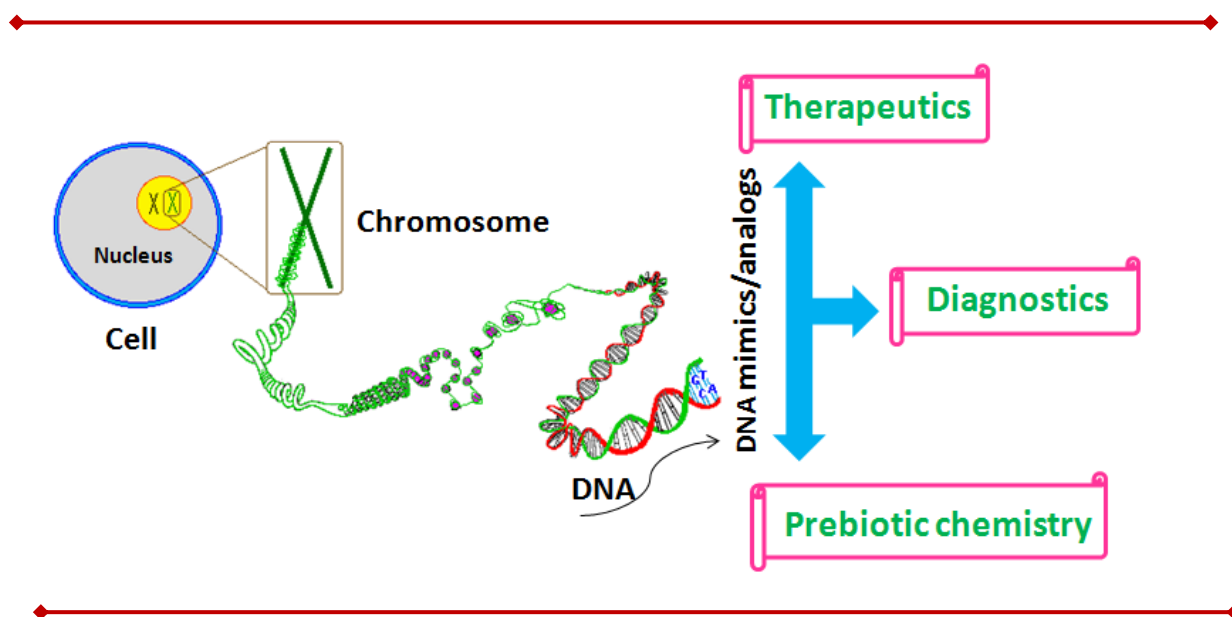
Table 7 Modified DNA sequences, their MALDI-TOF mass analyses

Sequences/Code	Mass	
	Calcd.	Obsd.
cctcttacctcagt $\underline{\mathbf{T}}^{\text{Ery}}$ aca ON1	5352.9	5116.2
cctcttacc $\underline{\mathbf{T}}^{\text{Ery}}$ cagt $\underline{\mathbf{T}}^{\text{Ery}}$ aca ON2	5338.8	5116.4
cttcttcc $\underline{\mathbf{T}}^{\text{Ery}}$ t ON3	2904.4	2616.5
cttct $\underline{\mathbf{T}}^{\text{Ery}}$ cc $\underline{\mathbf{T}}^{\text{Ery}}$ t ON4	2890.4	2615.9

In summary, we have designed an erythrulose-based oligonucleotides, from a four-carbon ketose L-erythrulose, whose abiotic formation was experimentally proved in the context of ‘glyoxylate scenario’. We have successfully achieved the monomer synthesis from a four-carbon tartaric acid, whose abiotic formation was also proved recently. Unfortunately the MALDI-TOF mass of the modified oligonucleotides did not match with the calculated, probably suggesting the need for the carbonyl protections of our designed monomer before its incorporation on DNA synthesizer. Additionally, like the 2’-OMe RNA analogue, we can also use OMe-substitution instead of having carbonyl in our design, which is exactly an acyclic version of the 2’-OMe RNA having backbone shortened by one-atom.

CHAPTER 1

Introduction to nucleic acids



1A Nucleic Acids: An introduction

1A.1 Chemical components and primary structure of nucleic acids

The realization that DNA was indeed the carrier of hereditary information led the scientists to understand how a molecule written only with the four chemical alphabets adenine (A), cytosine (C), guanine (G) and thymine (T) could encode the complex information that instructs the generation and functioning of all living beings.

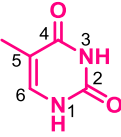
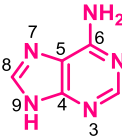
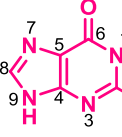
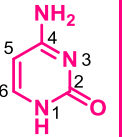
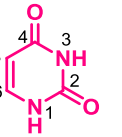
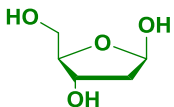
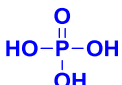
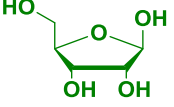
	Only in DNA	Both in DNA and RNA		Only in RNA	
Nitrogenous bases (Purine & Pyrimidine)	 <p>Thymine</p>	 <p>Adenine</p>	 <p>Guanine</p>	 <p>Cytosine</p>	 <p>Uracil</p>
Sugar and phosphate	 <p>2-Deoxyribose</p>	 <p>Phosphoric acid</p>	 <p>Ribose</p>		

Figure 1 Chemical components of nucleic acids.

Deoxyribose or Ribose nucleic acids (DNA or RNA) are polymers that are made up of nucleoside units joined sequentially via phosphodiester linkages. Each nucleoside consists of two components: a nitrogenous heterocyclic base, which is either a purine or a pyrimidine and a pentose sugar (deoxyribose or ribose in DNA or RNA respectively, Figure 1). The nitrogenous bases adenine, cytosine, and guanine are found in both RNA and DNA, while thymine only occurs in DNA and uracil only occurs in RNA (Figure 1). A very rare exception is a bacterial virus called PBS1 that contains uracil in its DNA.¹ Polymerization of the above mentioned relatively simple components leads to the DNA formation, as outlined in Figure 2. A deoxyribose sugar attached with the nucleobase is called as a nucleoside, a nucleoside phosphorylated at 5'-hydroxyl group is called as a nucleotide monomeric unit and the polymeric chain of these nucleotides is called as DNA. Anhydride-

like di- and tri-phosphate nucleotides have been identified as important energy carriers in biochemical reactions, the most common being adenosine 5'-triphosphate (ATP).²

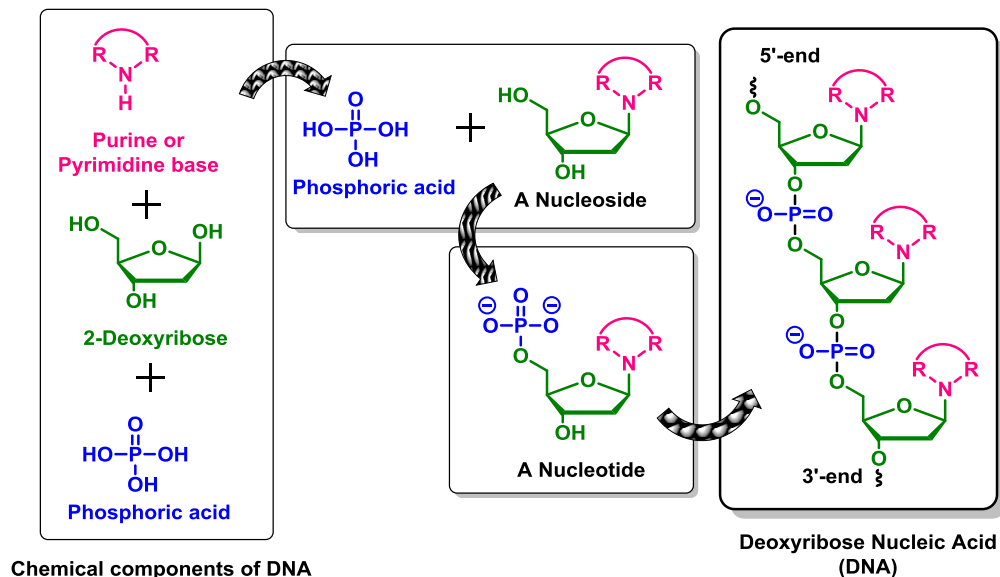


Figure 2 Sketch of chemical components leading to genetic material DNA.

1A.2 Base Pairing *via* Hydrogen bonding

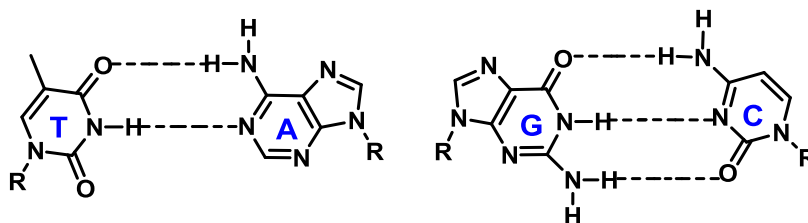


Figure 3 Watson and Crick base pairing of A: T and G: C

The N-H groups of the nucleobases are potent hydrogen bond donors, while the sp²-hybridized electron pairs on the oxygen of the base C=O groups and that on the ring nitrogens are hydrogen bond acceptors. In Watson-Crick base-pairing, there are two hydrogen bonds in an A:T base pair and three in a C:G base pair (Figure 3). Though the Watson-Crick base pairing is dominant between the nucleobases, other significant pairings are Hoogsteen (HG)³ and Wobble base pairs.⁴

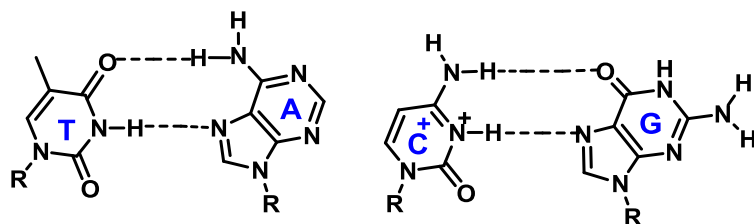


Figure 4 Hoogsteen base pairing of A: T and G: C.

A Hoogsteen A:T base pair (Figure 4) applies the N7 position of the purine base (as a hydrogen bond acceptor) and C6 amino group (as a donor), which bind the Watson-Crick (N3-O4) face of the pyrimidine base. Hoogsteen pairs have quite different properties from Watson-Crick base pairs. The angle between the two glycosylic bonds (ca. 80° in the A:T pair) is larger and the C1'–C1' distance (ca. 8.6 Å) is smaller than in the regular geometry. In some cases, called reversed Hoogsteen base pairs, one base is rotated 180° with respect to the other. Hoogsteen base-pairing allows sequence specific binding of pyrimidine third strands in the major groove of Watson-Crick purine: pyrimidine duplexes to form triple-helical structures (poly(dA):2poly(dT)) and (poly(rG):2poly(rC)).

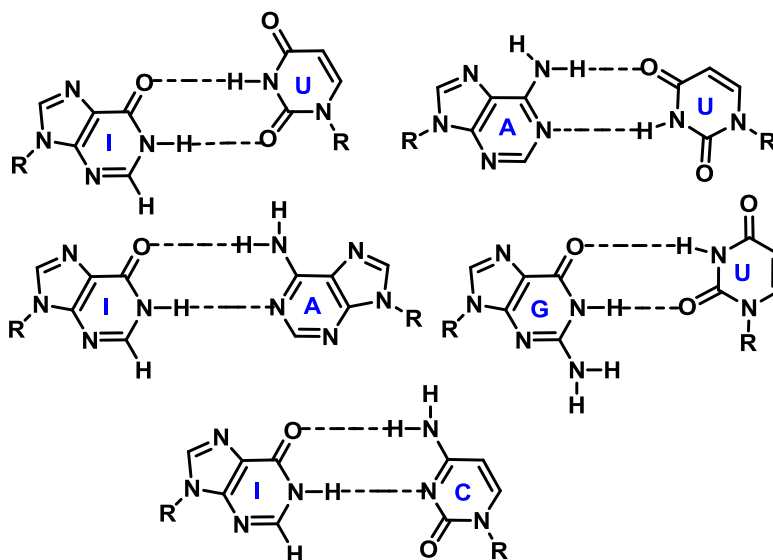


Figure 5 Wobble base pairing of Ionosine and Uracil.

In the wobble base pairing (Figure 5), a single purine base is able to recognize pyrimidines (e.g. G:U, where U= uracil) and have importance in the interaction of

messenger RNA (m-RNA) with transfer RNA (t-RNA) on the ribosome during protein synthesis (codon-anticodon interactions). Several mismatched base pairs and anomalous hydrogen bonding patterns have been seen in X-ray studies of synthetic oligodeoxynucleotides.⁵

1A.3 Sugar puckering in nucleos(t)ides

The molecular geometry of an individual nucleotide is very closely related to that of the corresponding nucleotide units in oligomers and nucleic acid helical structure. The details of the conformational structure of nucleotides are accurately defined by the torsional angles α , β , γ , δ , ϵ and ζ in the sugar-phosphate backbone, θ_0 to θ_4 , in the furanose ring, and χ for the glycosidic bond (Figure 6). Because many of these torsional angles are interdependent, one can simply describe the shapes of nucleotides in terms of the parameters: the sugar pucker, the syn-anti conformation of the glycosidic bond, the orientation of C4'-C5' bond and the phosphodiester backbone. The pentose sugar rings in nucleosides are twisted or puckered in order to minimize non-bonded stereo-electronic interactions between their substituents.

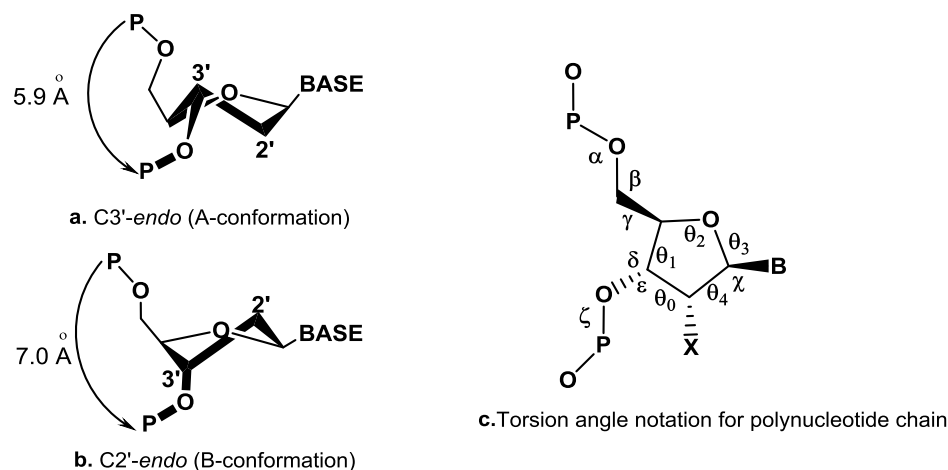


Figure 6 a) Structures of C3'-endo, b) C2'-endo preferred sugar puckers and c) Torsion angle notation for polynucleotide chain.

This 'puckering' is described by identifying the major displacement of carbon C-2' and C-3' from the median plane of C1'-O4'-C4'. Thus, if the *endo*-displacement of C-2' is

greater than the *exo*-displacement of C-3', the conformation is called C2'-*endo* and so on for other atoms of the ring (Figure 6a and 6b). The *endo*-face of the furanose is on the same side as C5' and the base; the *exo*-face is on the opposite face to the base. The sugar puckers are located in the north (N) and south (S) domains of the pseudorotation cycle of the furanose ring.⁶ In solution, N and S conformations are in rapid equilibrium and are separated by low energy barrier. The average position of the N \leftrightarrow S pseudorotational equilibrium is influenced by various steric and stereoelectronic effects (*gauche* and anomeric effect) as exemplified below with the nucleoside units.

A.3.1 Steric effect

On the basis of steric effect alone, for β -D-nucleosides, S-type pseudorotamers are energetically favoured in comparison with N-type counterparts, since the pseudoequatorially oriented nucleobase in the former exerts less steric repulsions with the other substituents on the pentofuranose moiety than when it is pseudoaxial in the later (Figure 7). Hence, as the bulk of the nucleobase is increased, it is expected to shift the two-state N \leftrightarrow S pseudorotational equilibrium towards more S-type conformation in β -nucleosides.

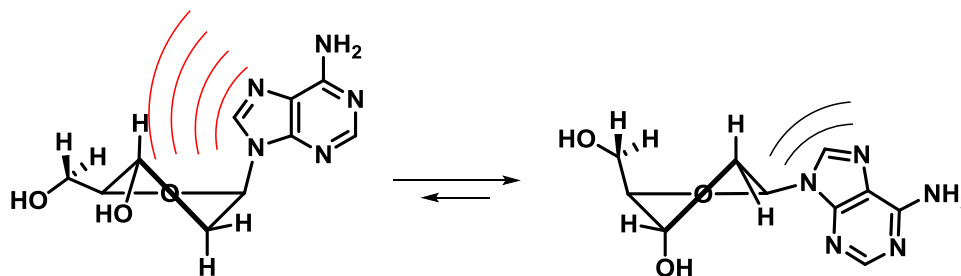


Figure 7 The drive of the N \leftrightarrow S pseudorotational equilibrium towards N-type over S-type by the steric effect of nucleobases, exemplified in β -D-2'-deoxyadenosine.

A.3.2 Anomeric effect (Edward-Lemieux effect)

The generalized anomeric effect in nucleosides/nucleotides results from an extension of the original Edward-Lemieux effect to describe the preference for *gauche* over *trans* orientation of R-Y with respect to R'-X in R-Y-R'-X fragments, where R-Y-R'-X

may be either acyclic or constitute a part of a ring, Y is a heteroatom with one or more lone pairs, and X is an electronegative atom or group.⁷

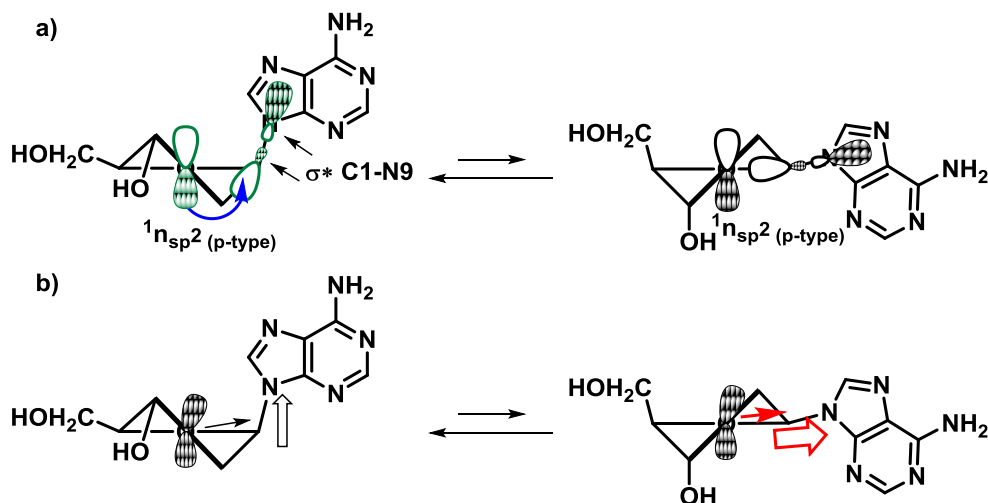


Figure 8 Anomeric effect supported by hyperconjugation model (a) and dipole-dipole repulsions model, exemplified in β -D-2'-deoxyadenosine (b).

The O4'-C1'-N1/N9 anomeric effect in nucleosides and nucleotides may be explained either (i) by stabilizing $n_{O4'} \rightarrow \sigma^*_{C1'-N1/N9}$ orbital interactions between the orbital of one of the endocyclic O4' electron lonepairs ($n_{O4'}$) and the antibonding orbital of the C1'-N1/N9 glycosyl bond ($\sigma^*_{C1'-N1/N9}$) (Figure 8a), or (ii) by destabilizing electrostatic repulsions in S-type pseudorotamer between the dipoles of inherent furanose ring (thick arrow in Figure 8b, resultant of C4'-O4', O4'-C1' individual dipole moments and of the dipole induced by O4' lonepairs) and the dipole oriented from C1' to N1/N9 (hollow arrow, Figure 8b).⁷ The N-type pseudorotamers of β -D-nucleosides shows the reduced dipole-dipole repulsions and enhanced O4'-C1'-N1/N9 hyperconjugative interactions than the S-type, hence driving $N \leftrightarrow S$ pseudorotational equilibrium towards N-type conformation.

The pyrimidine nucleosides (connected to anomeric carbon through -N1 of electron poor pyrimidine) exhibit stronger anomeric effect than the purine nucleosides (connected to anomeric carbon through -N9 of electron rich imidazole ring), as the antibonding orbital of former is having low energy and hence is a better acceptor than that of the latter. Similarly,

the nucleosides containing protonated nucleobases also have strengthened anomeric effect credited to the effective $n_{O4'} \rightarrow \sigma^*_{C1'-N1/N9}$ interactions.

A.3.3 Gauche effect

The term "gauche" refers to specific conformational isomers where two vicinal groups are separated by a 60° torsion angle, i.e. the "synclinal alignment of groups attached to adjacent atoms" (Figure 9a). The hyperconjugation offers a better explanation to the observed gauche effect. In the hyperconjugation model, the donation of electron density from σ bonding orbital of the best donor to the σ^* antibonding orbital of the best acceptor is considered the source of stabilization in the gauche isomer. Only the gauche conformation allows good overlap between the better donor and the better acceptor. In addition, the gauche effect implies the electronic energy preference of a gauche-rotamer over the corresponding anti-rotamer. Hence, the "gauche effect" can be generalized as the tendency of R-X-Y-R' and X-C-C-Y fragments (where X and Y represent electronegative atoms or groups) to adopt preferentially gauche over antiperiplanar orientations in spite of unfavourable steric repulsions and electrostatic interactions.

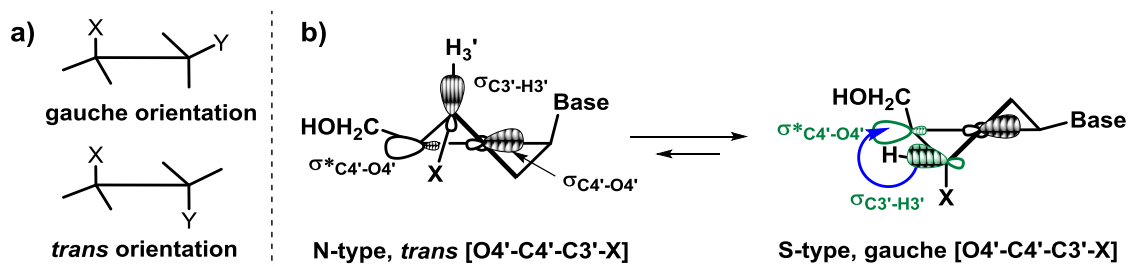


Figure 9 Gauche and *trans* orientation of X-C-C-Y fragment (a) and rationalization of O4'-C4'-C3'-X gauche effect of 2'-3'-dideoxy-3'-substituted β -D-nucleoside (b).

In β -2'-deoxynucleosides, the gauche effect of [O4'-C4'-C3'-O3'] stabilizes the S-type pseudorotamer in which the O4'-C4' and O3'-C3' bonds are in a gauche orientation (Figure 9b), over N-type (trans orientation of O4'-C4' and O3'-C3'), contrast to the anomeric effect which drives the N \leftrightarrow S pseudorotational equilibrium in solution toward N-type forms. The experimentally observed preference of β -2'-deoxy mono- and oligonucleotides for S-type conformations therefore suggested that the [O4'-C4'-C3'-O3']

gauche effect is the predominant factor in controlling the sugar conformation than the counteracting anomeric effect.⁷

In β -D-ribonucleosides compared with β -D-2'-deoxyribonucleosides, two additional gauche effects operate within [HO2'-C2'-C1'-O4'] and [HO2'-C2'-C1'-N1/N9] fragments. However, one of the additional gauche effect caused by the [HO2'-C2'-C1'-O4'] fragment was nullified by the gauche effect resulting from [HO3'-C3'-C4'-O4'] fragment. In spite of the cancellation of these gauche effects, the fine balancing of nucleobases and [HO2'-C2'-C1'-N1/N9] gauche effect drives the N \leftrightarrow S pseudorotational equilibrium to N-type over S-type conformation.⁷

1A.4 Structures of Nucleic acids

1A.4.1 The duplex structure

Three major DNA conformations are believed to be found in nature, A-DNA, B-DNA, and Z-DNA (Figure 10). The "B" form described by James D. Watson and Francis Crick is believed to predominate in cells.⁸ B-DNA is a right-handed double helix with a wide major-groove and a narrow minor-groove where the bases are perpendicular to the helical axis. It is 23.7 Å wide and extends 34 Å per 10 bp of sequence. A-DNA and Z-DNA differ significantly in their geometry and dimensions from B-DNA, although still form helical structures. At low humidity and high salt, the favored form is highly crystalline A-DNA while at high humidity and low salt, the dominant structure is B-DNA. In both A and B forms of DNA the Watson-Crick base pairing is maintained by *anti* glycosidic conformation of the nucleobases. The sugar conformation however, is different in both forms with the B form showing C2'-*endo* puckered sugar and the A form DNA exhibiting C3'-*endo* sugar-pucker. A very unusual form of DNA-duplex is the left-handed Z-DNA.⁹ This conformation of DNA is stabilized by high concentrations of MgCl₂, NaCl and ethanol, and is favored for alternating G:C/G:C sequences. Z-DNA has characteristic *zig-zag* phosphate backbone and the Watson-Crick base pairing is achieved by purines adopting *syn* glycosidic conformation with C3'-*endo* sugar-pucker.⁸ Some other possible DNA conformations are C-DNA, D-DNA, E-DNA, L-DNA, P-DNA and S-DNA.¹⁰

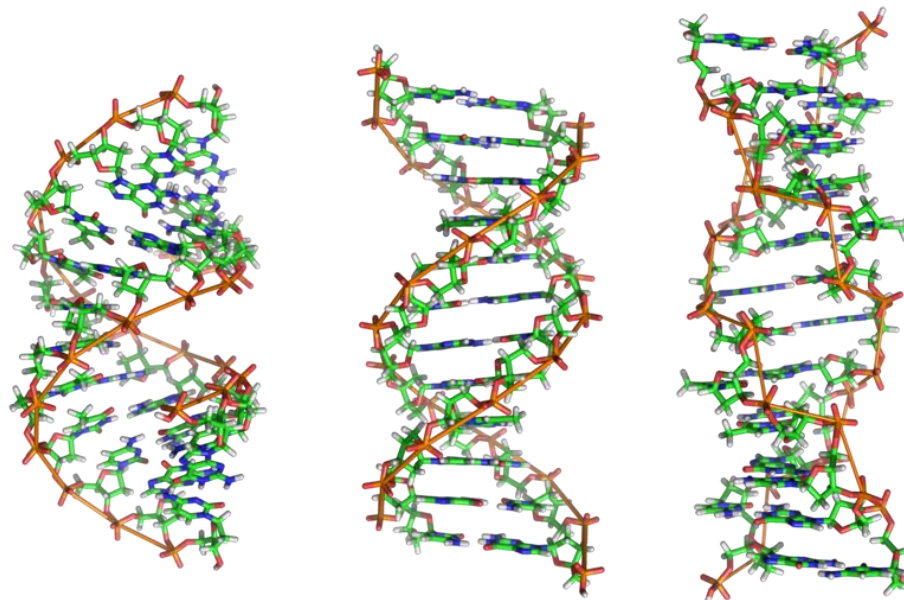


Figure 10 The structures of A, B and Z DNA in respective order (adopted from *wikipedia*).

RNA is recognized to have greater structural versatility than DNA.¹¹ RNA can form double stranded duplexes. The most important structural feature of RNA that distinguishes it from DNA is the presence of a hydroxyl group at the 2'-position of the ribose sugar, capable of intramolecular nucleophilic attack on adjacent phosphate. The presence of this functional group enforces the C3'-*endo* sugar conformation (as opposed to the C2'-*endo* conformation of the deoxyribose sugar in DNA) that causes the helix to adopt the A-form geometry and hinders the formation of B-type helix. This results in a very deep and narrow major groove and a shallow and wide minor groove.

RNA is single stranded but can form complex and unusual structures such as stem and bubble structures, due to the chain folding as a consequence of intramolecular base pairing. An example of folded RNA structure is *t*-RNA, which is the key RNA involved in the translation of genetic information from *m*-RNA to proteins. *t*-RNA contains about 70 bases that are folded such that there are base paired stems, bulges and open loops. The overall shape of the completely folded *t*-RNA is L-shaped. At low ionic strength, A-RNA has 11 base pairs per turn in a right-handed, antiparallel double helix. The sugars adopt a C3'-*endo* pucker and other geometric parameters are all very similar to A-DNA. If the salt

concentration is raised above 20%, A'-RNA form is observed which has 12 base pairs per turn of the duplex. Both structures have typical Watson-Crick base pairs, which are displaced by 4.4 Å from the helix axis, hence forming a very deep major-groove and rather shallow minor-groove.

The presence/absence of 2'-OH functional group in RNA/DNA controls the pseudorotational equilibrium of C3'-endo (compact backbone)↔C2'-endo (extended backbone). This kind of sugar puckering mainly affects the backbone torsion angle δ , which on average is $\sim 80^\circ$ in A-form duplex (RNA:RNA or RNA:DNA) and $\sim 120^\circ$ in B-form duplex (DNA:DNA).¹² The DNA:DNA duplex (C2'-endo conformation for every sugar unit) is an extended B-form structure whereas RNA:RNA duplex (C3'-endo conformation for every sugar unit) remains in compact A-form. As the conformational energy barriers in the 2'-deoxyribose sugar in DNA are relatively low than in the ribose sugar of RNA, the 2'-deoxyribose sugar has a flexibility to adopt C3'-endo (or O4'-endo) conformation in the DNA strands and allows DNA:RNA duplexes also to be in A-form. The C3'-endo conformation of ribose sugar in RNA, which has the *anti*-glycosidic nucleobases, projects the nucleobases in axial orientation and assists efficient base-pair stacking and hence strengthens the duplex stability with RNA/DNA. Based on the base-pairing strength, which is further enhanced by axially oriented nucleobases (in C3'-endo conformation), the duplex stability of natural nucleic acids (DNA/RNA) is in the order: RNA:RNA >RNA:DNA >DNA:DNA.

1A.4.2 Triplex-Forming Oligonucleotides (TFO)

Triple-helical nucleic acid structures are known since 1957.¹³ A purine base can form hydrogen-bonded contacts on two sides, one on Watson-Crick face and one on the Hoogsteen face. Thus, in duplex DNA, a purine stretch presents sites in the major groove for Hoogsteen complexation by a third strand.¹⁴ In 1991 it was realized by two different research groups that the single-stranded DNA can also serve as a target for triple helix formation: a purine stretch can be bound on two sides by a molecule carrying both a Watson-Crick complementary domain and a Hoogsteen complementary domain.

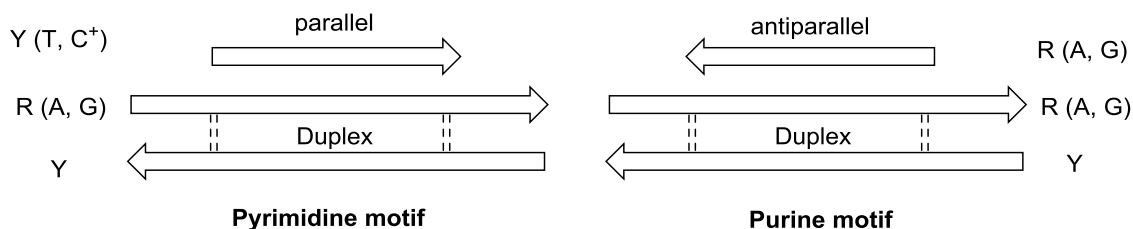


Figure 11 Pyrimidine and purine motifs of triplex formation.

The triplexes formed with synthetic oligonucleotides, remained an obscure part of DNA chemistry until 1987 when it was realized that they offered a means for designing sequence specific DNA targeting agents.^{14b} DNA triple helix formation results from the major groove binding of a third strand that is either pyrimidine (Y)- or purine-rich (R), in parallel (p) or antiparallel (ap) orientation respectively to the central purine strand as shown in Figure 11.¹⁵ A purine-rich third strand binds in antiparallel orientation to the central strand, while a pyrimidine-rich strand does so in a parallel orientation. The specificity in triplex formation is derived from Hoogsteen (HG) hydrogen bonding.

Thus, T recognizes A of A-T Watson-Crick base pair to form TpAT triad and protonated C binds to G of G-C base pair to give CH⁺ pGC in the pyrimidines motif (Figure 12). Similarly in the purine motif, the third strand A binds to A of A-T, leading to a AapAT triad, while G binds to G of G-C base pair forming GapGC triad in the reverse Hoogsteen mode.¹⁶

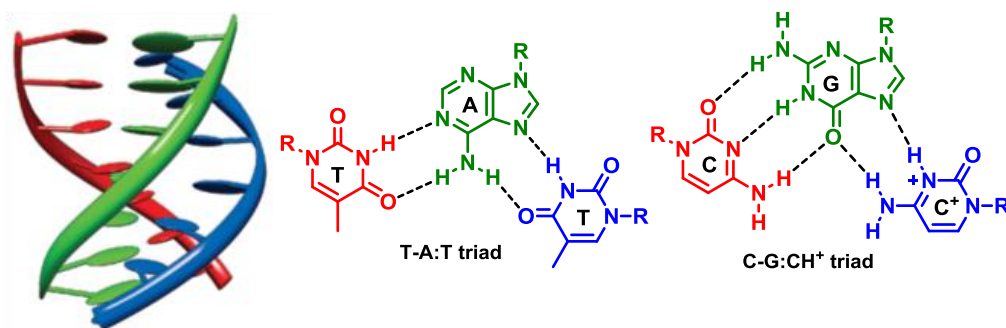


Figure 12 Schematic illustration of the parallel triplex and the base pairings involved (Figure adopted from ref.¹⁷).

For a given polypurine:polypyrimidine sequence of DNA it is possible to design therapeutic TFO that will specifically bind to it and thereby inhibit the gene expression. Different aspects of triple-stranded structures have been discussed in several reviews.¹⁸ The formation and stabilization of triplexes depends on different types of interactions like electrostatic forces, stacking and hydrophobicity contributions, Hoogsteen hydrogen bonds, and hydration forces. The triplex structure of single DNA molecule was hypothesized as an intermediate in the folding process of some quadruplex forming sequences.¹⁹ Limongelly *et. al* recently isolated the G-triplex of thrombin binding aptamer (TBA) and structurally characterized.¹⁷

1A.4.3 Quadruplex structure – The ‘G-quadruplex’

Telomeres, specific sequences present at both the ends of the eukaryotic chromosome, are essential for genome integrity and appear to play an important role in cellular aging and cancer.²⁰ Most of telomeric DNA is double-stranded except for the extreme terminal part where the 3' region of the G-rich strand is single-stranded.²¹ Both G-rich and C-rich telomeric strands may form *in vitro* unusual DNA structures. The G-rich strand can adopt a four-stranded G-quadruplex structure involving planar G-quartets,²² while the C-rich strand can form the so-called i-motif with intercalated C⁺C⁺ base pairs.²³

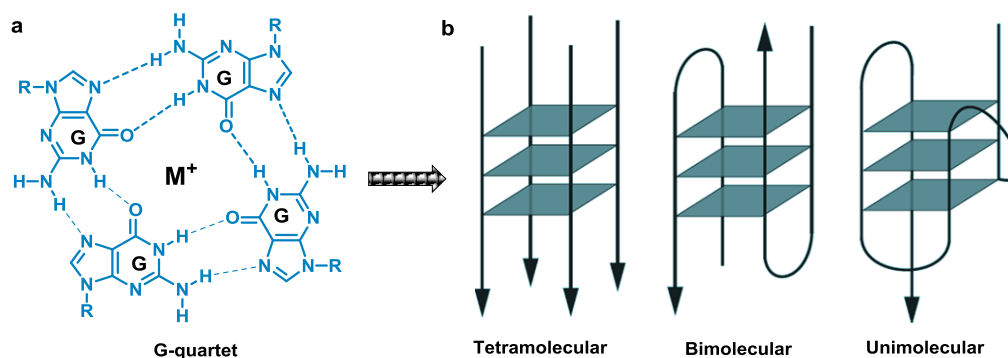


Figure 13 The G-quartet (a) and three G-quartets forming G-quadruplex (b) in tetramolecular/bimolecular/unimolecular fashion.

Guanine-quadruplexes (G-quadruplexes) are four stranded structures based on the stacking of planar guanine-quartets. The central cavity of G-quadruplexes is occupied by

cations which neutralize the electrostatic repulsion between guanine 'O6' oxygens and thus stabilize the overall structure. Quadruplexes can be formed with one (unimolecular), two (bimolecular) or four (tetramolecular) G-rich strands (Figure 13). Tetramolecular quadruplexes generally adopt a well-defined structure, in which all guanines are in the anti-glycosidic conformation and all strands are parallel, and might be useful for biotechnology applications.²⁴

1A.5 Spectroscopic technique to study DNA/RNA interactions

Ultraviolet absorption (UV) and circular dichroism (CD) which probe electronic properties of the bases are highly useful and general tools for characterizing nucleic acids.

1A.5.1 Ultraviolet spectroscopy

The absorbance of polynucleotide depends on the sum of the absorbance of the nucleotide plus the effect of the interaction among the nucleotides. The interaction cause a single strand to absorb less than the sum of its nucleotides and a double strand absorbs less than its two component single strand. The effect is called hypochromicity which results from the coupling of the transition dipoles between neighboring stacked bases and is larger in amplitude for A-U and A-T pairs.²⁵ In converse term, hyperchromicity refers to the increase in absorption when a double stranded nucleic acid is dissociated into single strands. The UV absorption of a DNA duplex increases typically by 20-30 % when it is denatured. This transition from a stacked, hydrogen bonded double helix to an unstacked, strand-separated coil has a strong entropic component and is temperature dependent. The mid-point of this thermal transition is known as the melting temperature (T_m). Such thermal dissociation of nucleic acid double helices to give single stranded DNA/RNA and the T_m is a function of base composition, sequence, chain length, salt concentration, and pH of the solvent (buffer). In particular, early observations of the relationship between T_m and base composition for different DNAs showed that A-T pairs are less stable than G-C pairs, a fact which is now expressed in a linear correlation between T_m and the gross composition of a DNA oligomer by the equation:

$$T_m = X + 0.41 (\% C + G) ^\circ C$$

The constant X is dependent on salt concentration and pH and has a value of 69.3 °C for 0.3 M sodium ions at pH 7.²⁵

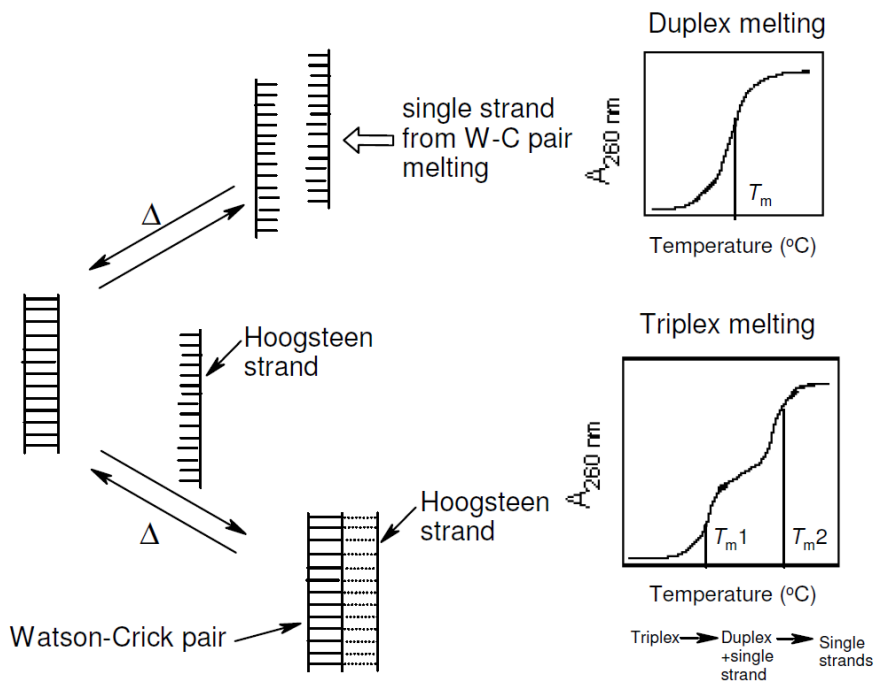


Figure 14 Schematic representation of duplex and triplex melting transitions

Duplex melting: According to the ‘all or none model’,²⁶ the UV absorbance value at any given temperature is an average of the absorbance of duplex and single strands. A plot of absorbance against temperature gives a sigmoidal curve in case of duplexes and the midpoint of the sigmoidal curve (Figure 14) called as the ‘melting temperature’ (T_m) (equilibrium point) at which the duplex and the single strands exist in equal proportions.

Triplex melting: In the case of triplexes, the first dissociation leads to melting of triplex generating the duplex (WC duplex) and third strand (Hoogsteen strand), followed by the duplex dissociation to form two single strands. The DNA triplex melting shows characteristic double sigmoidal transition (Figure 14) and UV melting temperature for each transition is obtained from the first derivative plots. The lower melting temperature (T_{m1}) corresponds to triplex to duplex transition while the second higher melting temperature (T_{m2}) corresponds to the transition of duplex to single strands.²⁷

1A.5.2 Circular Dichroism

Circular dichroism (CD) refers to the differential absorption of left and right circularly polarized light. CD is particularly useful for studying chiral molecules and has very special significance in the characterization of biomolecules (including the secondary structure of proteins and the handedness of DNA). The commonly used units in current literature are the mean residue ellipticity ($\text{degree cm}^2 \text{dmol}^{-1}$) and the difference in molar extinction coefficients called as molar circular dichroism or $\Delta\epsilon$ ($\text{liter mol}^{-1}\text{cm}^{-1}$). The molar ellipticity $[\theta]$ is related to the difference in molar extinction coefficients by $[\theta] = 3298.2 (\Delta\epsilon)$. In the nucleic acids, the heterocyclic bases that are principal chromophores. As these bases are planar, they don't have any intrinsic CD. CD arises from the asymmetry induced by linked sugar group. CD spectra from the dinucleotide exhibit hyperchromicity, being more intense by roughly an order of magnitude than those from monomers. It is when the bases are linked together in polynucleotide's, giving rise to many degenerate interactions and they gain additional characteristic associated with the asymmetric feature of secondary structure such as in proteins and nucleic acids.²⁸

The simplest application of CD to DNA structure determination is for identification of polymorph present in the sample.²⁹ The CD signature of the *B*-form DNA as read from longer to shorter wavelength is a positive band centered at 275 nm, a negative band at 240 nm, with cross over around 258 nm. *A*-DNA is characterized by a positive CD band centered at 260 nm that is larger than the corresponding *B*-DNA band, a fairly intense negative band at 210 nm and a very intense positive band at 190 nm. The 250-230 nm region is also usually fairly flat though not necessarily zero.

1B RNA therapeutics through antisense mechanism

1B.1 Introduction

Traditional small molecule drugs target proteins of disease origin such as enzymes and receptors, which represent a small subset of total cellular proteins. In contrast, oligonucleotides target pre-mRNA or mRNA- the carriers of genetic information before it is translated into proteins. Because mRNAs encode all cellular proteins, oligonucleotides

targeting mRNA could prove to be effective as targets and diseases that are not treatable by current drugs. A plethora of synthetic DNA mimics/analogues are known in the literature, that are capable of forming strong duplex with the target complementary DNA/RNA and also conferring the necessary metabolic stability.

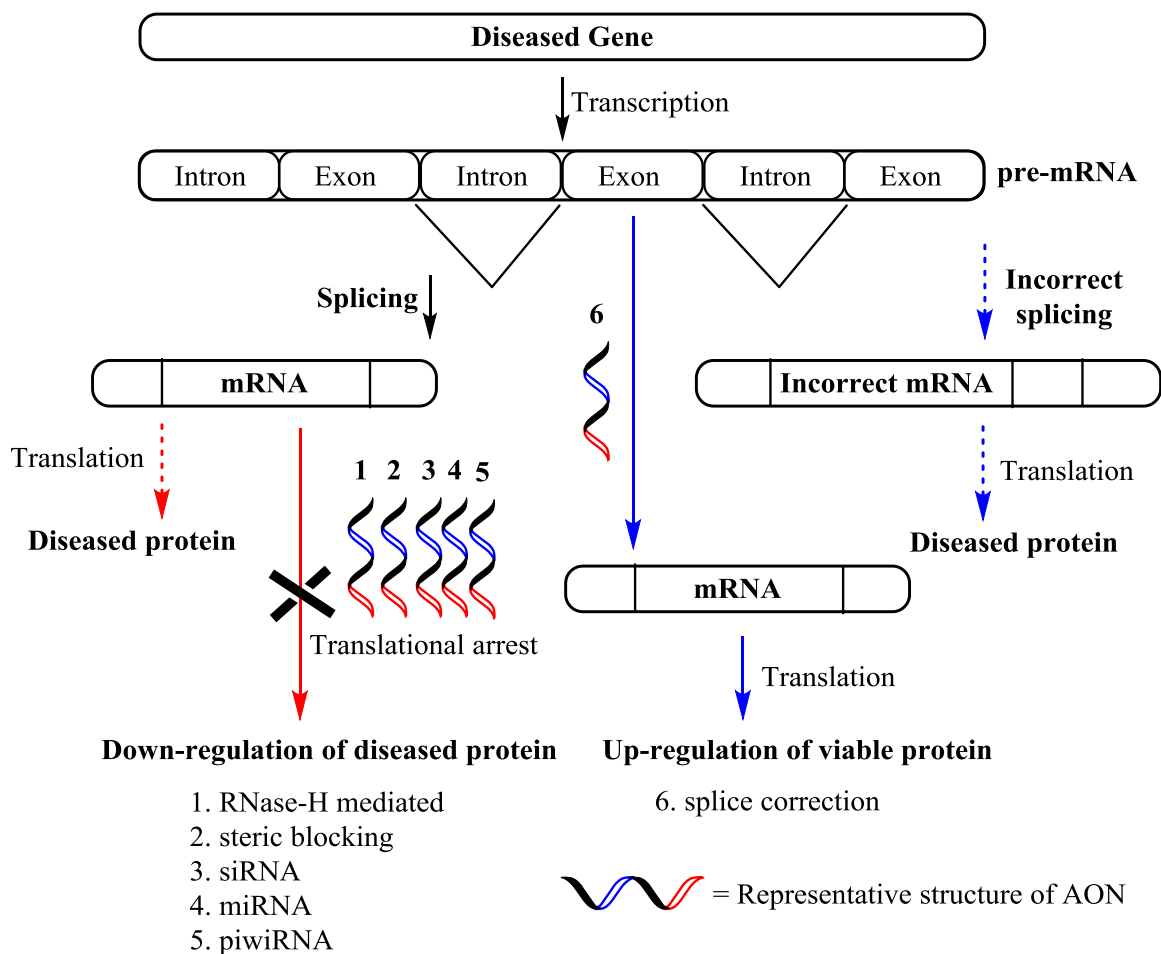


Figure 15 Sketch of protein synthesis modulation using antisense oligonucleotides (AONs).

Zamecnik and Stephenson's discovery of viral replication inhibition by the complementary 13-nucleotide-long DNA have witnessed the birth of the antisense technologies (Figure 15).³⁰ The RNA therapeutics target mRNA sequences through disruptive or corrective activity with strong and specific interaction of synthetic AON with target mRNA. The disruptive antisense down-regulates the synthesis of disease-causing

functional proteins by 'translational arrest' *via* **RNase-H activation**, or through **steric-blocking oligonucleotides**, or through RNAi - **RNA interference** (either small interfering RNA - siRNA or micro RNA - miRNA or piwi-interacting RNA – piRNA) (Figure 15). The corrective antisense approach up-regulates the gene expression of viable proteins by the **splice correction** of pre-mRNA (Figure 15). All these approaches involve the binding of complementary oligonucleotides/DNA analogues to target RNA through base pairing, and therefore all are operating by an antisense mechanism besides their substantial differentiation in mechanisms of action and the functional outcomes.

1B.2 RNase-H activation

The wealth of literature contains plethora of synthetic DNA mimics/analogues which are capable of forming strong duplex with the target complementary DNA/RNA and also conferring the necessary metabolic stability. Numerous chemical modifications that improve the drug-like properties of DNA have been evaluated for RNA therapeutics. Though U.S. Food and Drug Administration (FDA) approved two antisense drugs Fomivirsen- treatment for the cytomegalovirus retinitis³¹ (marketed as Vitravene in 1998, but was discontinued in 2004) and Mipomersen- treatment for homozygous familial hypercholesterolemia (marketed as Kynamro in 2013),³² several antisense drugs are under different phases of clinical trials.³³ Currently, a typical antisense oligonucleotide drug candidate is about 20 nucleotides in length with internucleosidic linkage as a phosphorothioate backbone. In addition, to increase the *in vivo* stability of antisense oligonucleotides both 5'- and 3'- ends were flanked with ~five modified nucleotides.

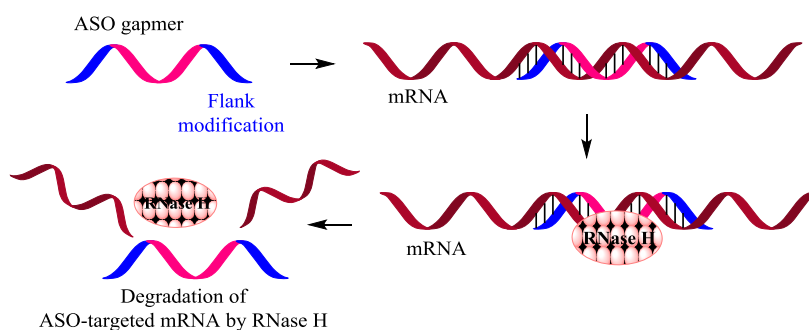


Figure 16 Schematic representation of degradation of mRNA by antisense oligonucleotides.

The remaining central 10-nucleotide phosphorothioate gap (hence the term ‘gapmers’) that allows the cleavage of targeted mRNA by ribonuclease-H (RNase-H), Figure 16. In contrast to siRNA, which tolerates only limited modifications to remain RISC-compatible, more extensive chemical modifications in flanking region of the phosphorothioate gapmers do not abrogate RNase H activity. One such modification, 2'-*O*-methoxyethyl (2'-MOE) ribonucleoside, is used in recently FDA approved antisense drug, Mipomersen. Mipomersen is a 2'-MOE phosphorothioate antisense oligonucleotide gapmer that targets mRNA encoding apolipoprotein B100 (APOB100) expressed in the liver; APOB100 is a protein that is involved in the production of low-density lipoprotein cholesterol (LDL-C; also known as ‘bad’ cholesterol). Mipomersen reduces the levels of APOB100 secreted from liver cells into the bloodstream.

1B.3 Steric-blocking oligonucleotides

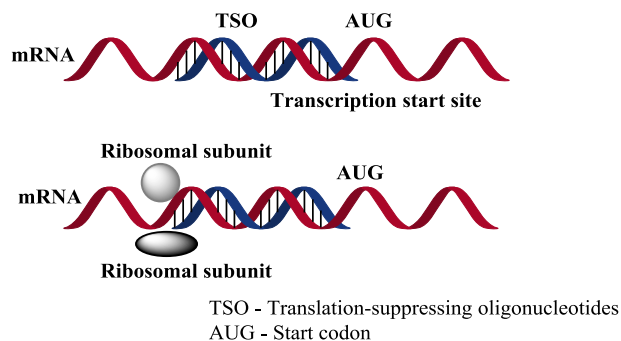


Figure 17 Down regulation of gene expression by the translation blocking of mRNA.

In parallel to the development of gapmer antisense oligonucleotides, oligonucleotides were developed that do not induce RNase-H mediated cleavage of mRNA but instead act by blocking targeted RNA without inducing its degradation (Figure 17). Early examples of these oligonucleotides comprise oligonucleoside methylphosphonates³⁴ and phosphordiamidate morpholino oligomers (PMOs).³⁵ The outcomes of this approach depend on the nucleotide sequence elements in mRNA or pre-mRNA that are targeted. These outcomes can include modulation of splicing when pre-mRNA is targeted, blocking of mRNA translation or RNA folding, and external guide sequence (EGS)-directed mRNA

degradation by a *t*RNA-processing ribozyme, RNase-P. They can also be used to block toxic RNAs that would otherwise seize protein factors at their expanded triplet repeats.

1B.4 RNA interference (RNAi)

RNAi was first demonstrated in *Caenorhabditis elegans*, in which the delivery of exogenous, long, double-stranded RNA (dsRNA) effectively silenced the expression of a gene by inducing the degradation of a homologous host mRNA.³⁶ The process of RNA interference can be moderated by either siRNA or miRNA, but there are subtle differences between the two. Both are processed inside the cell by the enzyme called Dicer and incorporated into a complex called RISC (RNA induced silencing complex). siRNA, however, is considered exogenous double-stranded RNA that is taken up by cells, or enters via vectors like viruses, while miRNA is single stranded and comes from endogenous (made inside the cell) non-coding RNA, found within the introns of larger RNA molecules. The other class of small RNAs used for RNA silencing is piwi-interacting RNA (piRNA).³⁷

1B.4.1 siRNA (small interfering RNA)

In 2001, Elbashir *et al.*³⁸ demonstrated that it was possible to silence gene expression by direct introduction of small interfering RNAs (siRNA) into mammalian cells, and hence avoid the Dicer catalysed reaction. siRNAs are the double-stranded RNA fragments that are 21–22 nucleotides long and interact with a multiprotein RNA-induced silencing complex (RISC). Within the RISC, the siRNA is unwound, the sense strand is discarded, and the antisense or guide strand binds to mRNA (Figure 18).³⁹ When siRNA is fully complementary to its target, the endonuclease argonaute 2 - a component of the RISC - cleaves the mRNA 10 and 11 nucleotides downwards from the 5' end of the antisense strand.⁴⁰ The catalytic component that cleaves the target RNA (slicer activity), has been identified as the protein designated Argonaut 2 (Ago2).⁴¹ Ago2 contains a domain which resembles RNase-H,⁴² a long-known protein that cleaves the RNA component of a DNA/RNA duplex. After cleavage, the target RNA lacks those elements which are typically responsible for stabilizing mRNAs, namely the 5' end cap and the poly-A tail at

the 3' end, so that the cleaved mRNA is rapidly degraded by RNases and the coded protein can no longer be synthesized.

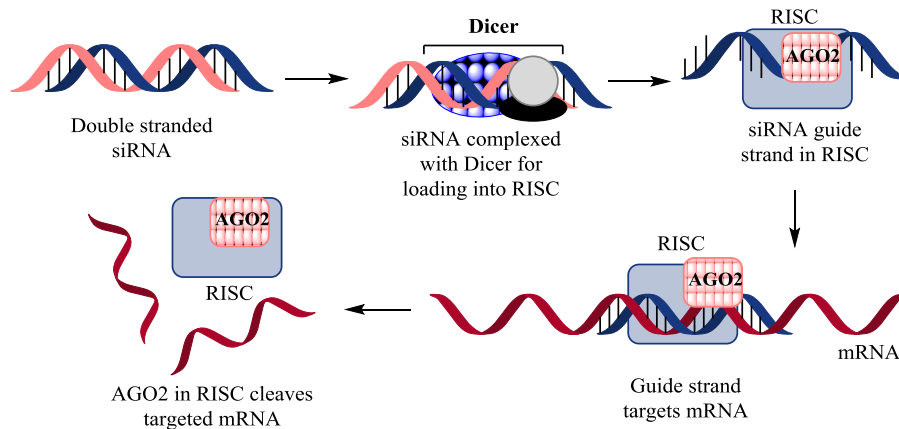


Figure 18 siRNA mediated down regulation of gene expression.

1B.4.2 miRNA (microRNA)

MicroRNAs (miRNAs),⁴³ are a class of non-protein-coding RNAs that regulate gene expression post-transcriptionally. They are single-stranded RNAs (ssRNAs) of ~19-25 nucleotides in length, generated from endogenous hairpin transcripts.⁴⁴ miRNAs can also function to silence mRNA much like short interfering RNAs (siRNAs) by integrating into the RISC and guiding the degradation of target message. miRNAs regulate the gene expression by base-pairing to partially complementary sites on the target messenger RNAs (mRNAs), usually in the 3' untranslated region (UTR). Binding of a miRNA to the target mRNA typically leads to translational repression and exonucleolytic mRNA decay, although highly complementary targets can be cleaved endonucleolytically. miRNAs have been found to regulate more than 30% of mRNAs and have roles in fundamental processes, such as development, differentiation, cell proliferation, apoptosis, and stress responses. Both loss and gain of miRNA function contribute to cancer development through a range of different mechanisms.⁴⁵ Because miRNAs regulate cancer cell differentiation, proliferation, survival, and metastasis, manipulating miRNA function, either by mimicking or inhibiting miRNAs implicated in cancer, could provide a powerful therapeutic strategy to interfere with cancer initiation and progression.

1B.4.3 piRNA (Piwi-interacting RNA)

piRNAs, a largest class of small non-coding RNAs, are 26–32 nucleotides in length and associate with Piwi proteins, such as Mili and Miwi and function in transposon silencing through heterochromatin formation or RNA destabilization *via* the formation of RISC.⁴⁶ They are distinct from microRNA in size (26–31 nt rather than 21–24 nt), lack of sequence conservation, and increased complexity.³⁷

1B.5 Splice-switching oligonucleotides

Precursor mRNA (pre-mRNA) is an incompletely processed single strand, synthesized from a DNA template in the nucleus of a cell by transcription. The completely processed pre-mRNA termed as the "mRNA". Pre-mRNA includes two different types of segments, exons and introns. Spliceosomes, comprised of small proteins and nuclear RNA, transcribe the exons (which will encode for protein) to mRNA, where as the non-protein coding introns were spliced out.⁴⁷ Oligonucleotide-induced modulation of splicing leads to several outcomes in cell culture and *in vivo* that have potential therapeutic value. These outcomes are not achievable with siRNA or classic gapmer antisense oligonucleotides, which only down-regulate gene expression.

Splice-switching oligonucleotides (SSOs) are the chemically modified oligonucleotides that block sequences in pre-mRNA splicing, and redirect mRNA splicing pathways. SSO–pre-mRNA duplexes are not recognized by RNase-H or the RISC, and the pre-mRNA is thus not cleaved. SSOs modulate pre-mRNA splicing, can repair defective RNA and restore the production of essential proteins. They can also generate novel proteins with desirable properties and regulate the presence of disease-related splice variant proteins, which can be achieved by modulation of alternative splicing of pre-mRNA.⁴⁸

1B.6 Alternative splicing

Alternative splicing is a regulated process during gene expression that results in multiple proteins synthesis from a single gene coding. During the process, particular exons of a gene may be included within, or excluded from, the final processed messenger RNA

(mRNA) produced from that gene.⁴⁹ Consequently the proteins translated from alternatively spliced mRNAs will contain differences in their amino acid sequence and their biological functions. This can be adapted to correct the alternative splicing process. So, the techniques that trick the splicing machinery to correct the alternative splicing pathways can be of high therapeutic value. There are numerous modes of alternative splicing observed, of which the most common is exon skipping. In this mode, a particular exon may be included in mRNAs under some conditions or in particular tissues, and omitted from the mRNA in others.⁴⁹ As over 95% of all human genes produce splice variant proteins by alternative splicing, modulation of alternative splicing may be applicable to a multitude of diseases.⁵⁰ Approximately 15% of all genetic diseases are caused by mutations that affect RNA splicing.

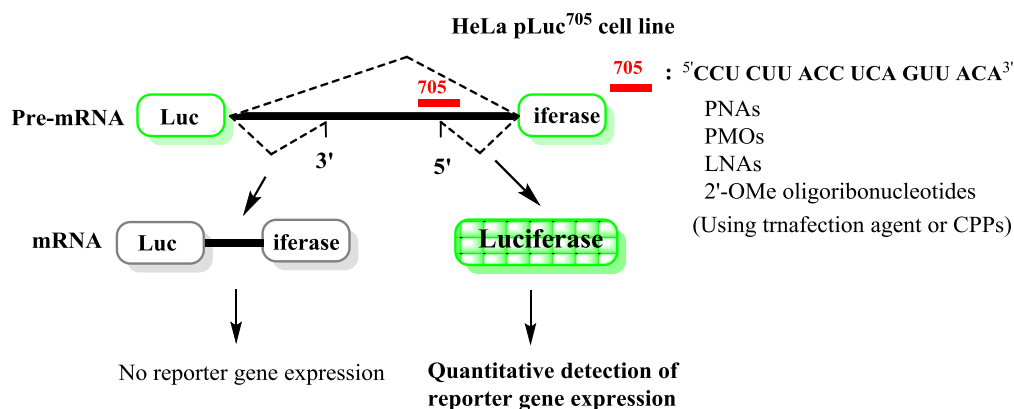


Figure 19 Splice correction leading to luciferase expression through a reporter gene.

To modulate pre-mRNA splicing, oligonucleotides must block RNA sequences that are essential for splicing and prevent the interaction of splicing factors - such as RNA-binding proteins, small nuclear RNAs and other components of the spliceosome with the pre-mRNA. The chemistries that have been shown to work in animal models include peptide nucleic acids (PNAs), alternating locked nucleic acids (LNAs) and deoxynucleotide oligonucleotides, fully modified (non-gapmer) 2'-substituted oligonucleotides and PMO-based oligomers.⁵¹ The latter two SSO chemistries have been used in clinical trials that tested splicing modulation as a treatment for DMD (Duchenne muscular dystrophy).⁵²

Kang and co-workers have developed an assay based on fluorescent protein expression as a result of steric blockage of alternative splicing by an antisense agent (synthetic mimic/analogue of DNA) (Figure 19).⁵³ This assay is straightforward to carry out and has a very high dynamic range, such that even very low activity levels can be seen as a positive luminescence read-out. The developed assay would be ideal even to assess the efficiency of cellular delivery of particular antisense oligonucleotides.

1B.7 Some promising chemical modifications of nucleic acids

DNA/RNA- natural nucleic acids, could not be used as oligonucleotide drugs because of their instability in biological systems and also their low binding ability to the target. To improve the potency of the oligonucleotides chemical modifications are necessary. There are plethora of modifications in the literature that were designed and evaluated for antisense therapeutics (as briefly drawn in Figure 20). This section offers the most promising modifications in the literature that are developed over the few decades. Among the earliest modifications the most successful include the phosphorothioates (PS) and phosphoramidate morpholino oligomers (PMOs).

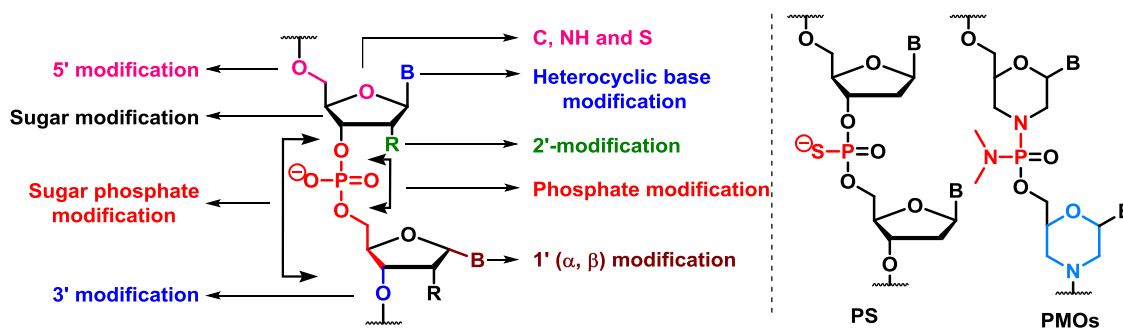


Figure 20 Schematic showing the extensive chemical modifications of DNA/RNA and the chemical structure of earliest modifications PS and PMOs.

A simple substitution of a *S*-atom for one of the non-bridging *O*-atoms of the internucleoside phosphate linkage led to the phosphorothioates (PS) (Figure 20), the first generation modifications.⁵⁴ Phosphorothioate-containing antisense oligonucleotides (ASOs) are able to activate RNase-H for target mRNA cleavage, while also having enhanced drug-like properties such as increased nuclease resistance, increased binding to cellular and

serum proteins. PMOs are the result of replacement of sugar of DNA by a morpholine ring and also the internucleosidic ionic phosphate linkage by a non-ionic phosphorodiamidate inter-subunit linkage (Figure 20). PMOs act by steric blocking to knock down the gene,⁵⁵ leading to the down-regulation of diseased protein synthesis, or they can up-regulate the viable protein synthesis by modifying the splicing of pre-mRNA.⁵⁶ However, first-generation antisense drugs also have limitations with regards to RNA binding strength, toxicology, pharmacokinetics, and pharmacodynamics.⁵⁷ This led to the further modifications to address these limitations.

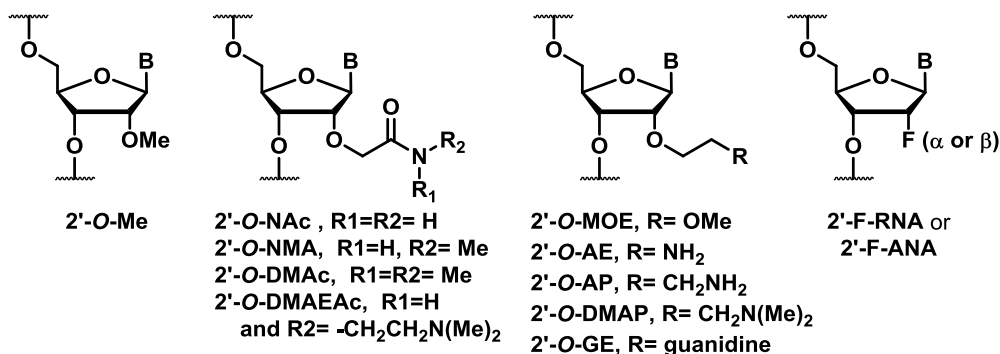


Figure 21 Few representative 2'-sugar modified oligonucleotides.

In DNA the *gauche effect* between O3' and O4' results in a C2' *-endo* or southern (S) sugar conformation. By modifying the sugar with electronegative substituents at the C2' and C3' position the overall conformation of the sugar (dictated by the net *gauche* and anomeric effect, as discussed in '*sugar puckering*' of this chapter) can be locked or frozen to yield oligonucleotides that bind target DNA/RNA with high affinity. In order to improve the RNA binding capacity of AON, second generation modifications focused mimicking RNA or C3'-*endo*-like structures. Electronegative substituent's such as fluorine and oxygen shift the ribose conformational equilibrium towards the C3'-*endo* pucker.⁵⁸ The 2'-fluoro substitution in both ribo-(α -F) and arabino-(β -F) configuration (Figure 21) led to stabilization of duplexes with target RNA, both operating in different mechanism. The observed duplex stability is attributed to the preferential C3'-*endo* sugar puckering of ribofluoro-modified (stabilized by both *anomeric effect* and *gauche effect*)⁵⁹ and the

favourable inter-residual pseudo-hydrogen bond (2'F...purine H8) in the arabinofluoro-modified nucleic acids.⁶⁰

The effect of 2'-substitution on the thermal stability of a DNA/RNA duplex was demonstrated with 2'-*O*-Me modification (Figure 21).⁶¹ The 2'-*O*-Me modification improved thermal stability of an oligonucleotide hybridized to a complementary RNA but did not confer the necessary metabolic stability to antisense oligonucleotides. In combination with phosphorothioate backbone modification, oligonucleotides with these modifications are resistant to metabolic degradation. Oligonucleotides with 2'-*O*-[2-(methoxy)ethyl] modification (2'-*O*-MOE, Figure 21) showed equal or higher binding affinity than 2'-*O*-Me modified oligonucleotides (attributed to the additional gauche effect of 2'-*O*-methoxyethyl tether),⁶² also offers +2° increase in thermal stability (T_m) per modification compared to PS-DNA. To improve the antisense properties of 2'-*O*-MOE-modified oligonucleotides, several novel 2'-*O*-modifications analogs of the parent have been synthesized (Figure 21).⁶³ These 2'-modified oligonucleotides showed excellent binding affinities to complementary RNA and high nuclease stability. Among these 2'-*O*-NMA⁶⁴ is a suitable modification for RNase-H based antisense therapeutics.⁶⁵

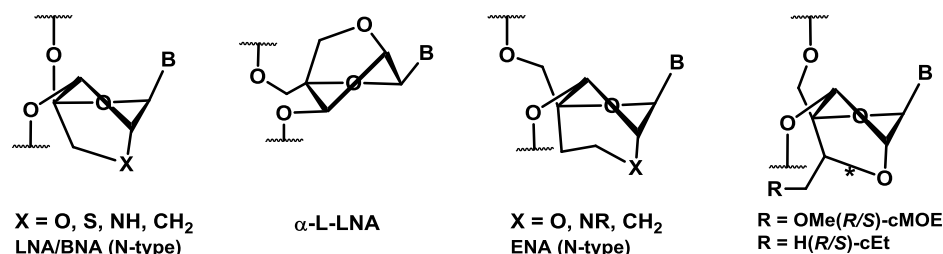


Figure 22 Locked nucleic acid and select analogs.

The electronegative substitutions at C2' causes the sugar to adopt the C3'-*endo* conformation, which leads to an increase in duplex stability, while locked nucleic acids (LNA) are a class of high-affinity RNA analogs in which the ribose ring itself is “locked” in C3'-*endo* conformation (Figure 22).⁶⁶ As a result, LNA oligonucleotides exhibit unprecedented thermal stability when hybridized to a complementary DNA or RNA strand. For each incorporated LNA monomer, the melting temperature (T_m) of the duplex increases

by 2-8 °C compared to the unmodified duplexes.⁶⁷ In addition, LNA oligonucleotides can be made shorter than traditional DNA or RNA oligonucleotides and still retain a high T_m . This is important when the oligonucleotide is used to detect small or highly similar targets.

In recent years, numerous LNA analogs have been reported. Replacing the 2'-O-atom in LNA with a N-/S-/C- atom provided, amino-LNA, thio-LNA and carba-LNA respectively (Figure 22) and the thermal stability of the LNA analogs with the target were evaluated. While amino-LNA decreased the thermal stability of the duplex,⁶⁸ thio-/carba-LNA retained the LNA duplex stability with the target RNA.⁶⁹ Koizumi *et al*⁷⁰ synthesized ethylene-bridged nucleic acid (ENA, Figure 22) with six-membered bridged structure. ENA demonstrated similar affinity and improved nuclease resistance, relative to LNA.^{69, 71} Replacement of the 2'-O-atom with a N-atom in the ENA gave AzaENA⁷² and with a C-atom gave carboENA. Both ENA analogs showed reduced thermal stability relative to ENA. Combining the structural elements of 2'-O-methoxyethyl (MOE) and LNA nucleosides as cMOE and cEt modifications of LNA retained the same affinity for the target, additionally enhanced the nuclease stability.⁷³

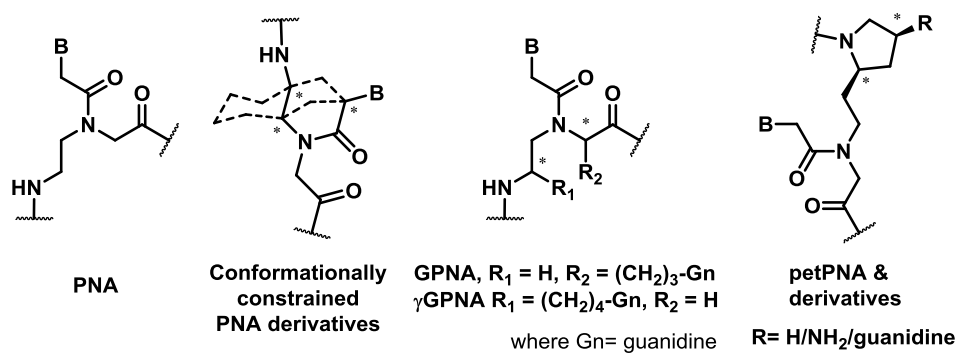


Figure 23 Peptide nucleic acid and select analogs.

Peptide nucleic acid (PNA, Figure 23) is acyclic, neutral and achiral DNA mimic that binds to complementary DNA/RNA sequence with high affinity and sequence specificity.⁷⁴ In PNA, the natural nucleobases are attached *via* methylene carbonyl linkers to an uncharged pseudopeptide backbone composed of repeating *N*-(2-aminoethyl)glycyl units. PNA hybridizes to complementary DNA/RNA sequences via specific base

complementation to form duplexes for mixed sequences and triplexes for homopyrimidine/homopurine sequences.⁷⁵ The main limitations of PNAs are its poor water solubility and lack of cell permeability coupled with ambiguity in DNA/RNA recognition arising from its equally facile binding in a parallel/antiparallel fashion with the target sequence. These limitations are being systematically addressed with rationally modified PNA analogues.⁷⁶

Chirality and the structural constraints as in conformationally constrained PNA derivatives (Figure 23) have been introduced in the backbone to address problems of equal binding of target in parallel/antiparallel fashion and the pre-organization (*i.e.* to enhance the target selectivity RNA vs DNA). To improve the aqueous solubility and cellular uptake, cationic residues were incorporated in the backbone through a linker, illustrated by GPNA (guanidine-based peptide nucleic acid)⁷⁷ and amino-/guanidino petPNA [*N*-(pyrrolydiny-2-ethyl) PNA] (Figure 23).⁷⁸ Due to the exceptional properties, PNA and analogues are exploited in wide range of applications such as in molecular biology,⁷⁹ developing lead compounds for gene targeted drugs via antisense/antigene technology,⁸⁰ biosensors⁸¹ and diagnostics.⁸²

1C Biomedical applications of synthetic DNA together with graphene/graphene oxide

1C.1 Introduction

As carbon is one of the most common elements in our ecosystem, carbon materials are known to be more environmentally and biologically friendly than inorganic materials. In particular, graphite- a naturally occurring material that has been used in our daily lives for hundreds of years without critical toxicity issues. Thus, one would expect that graphene, a novel single-atom-thick and two-dimensional carbon material with extraordinary electronic, mechanical and thermal properties,⁸³ would be also safe and useful for biological purposes. The carbon nanotubes (CNTs) which do not exist naturally, possessing extreme one-dimensional morphology found to be rather cytotoxic. In comparison, the two-dimensional shape of graphene is expected to be negligibly harmful in mild concentration,

so graphene materials would be readily applicable to biomedical research. Graphene is now expanding its territory beyond electronic and chemical applications toward biomedical areas such as precise biosensing through graphene- quenched fluorescence, graphene-enhanced cell differentiation and growth, and graphene-assisted laser desorption/ionization for mass spectrometry.⁸⁴

Despite its short history, the discovery of graphene has aroused large interest in the scientific community. Specifically, graphene oxide (GO), chemically exfoliated from oxidized graphite, which displays good water-solubility and flexibility for modification due to the massive number of suspended hydroxyl and carboxyl groups present on the surface,⁸⁵ has also attracted great interest in biological and biomedical application areas. In view of GO's properties of long-range nanoscale energy transfer and adsorption with nucleobases or aromatic compounds noncovalently via π -stacking interaction,⁸⁶ a number of sensors have been developed to detect nucleic acids,⁸⁷ metal ions,⁸⁸ enzyme activity,⁸⁹ and medicines.⁹⁰ Graphene-based materials possess large specific surface area, excellent conductivity, and availability for surface functionalization, which are important characteristics in the electrochemical applications.⁹¹ Recently, many graphene-based electrochemical sensors to detect glucose, ascorbic acid, dopamine, H₂O₂, DNA bases, and antigen has been reviewed.⁹²

1C.2 GO FRET Biosensor

The single-strand DNA (ssDNA) adsorb strongly with GO because the nucleobases are exposed and hence can interact with GO surface through strong π - stacking, where as the double-strand DNA (dsDNA) shows comparatively less interaction with the GO surface as the nucleobases are effectively hidden in duplex helical structure. Many GO-based biosensing platforms rely on GO's preferential interaction with ssDNA over dsDNA. When fluorescently labeled ssDNA probes adsorbed to the GO surface are hybridized with its complementary target ssDNA, are desorbed from GO by forming a DNA duplex, the fluorescence that was quenched by GO gets recovered.^{87a} Based on this concept, the detection of multiple ssDNA⁹³ and microribonucleic acid (miRNA) was successfully

demonstrated. The limit of detection (LOD) was further improved by cyclic enzymatic amplification, Figure 24.^{87d}

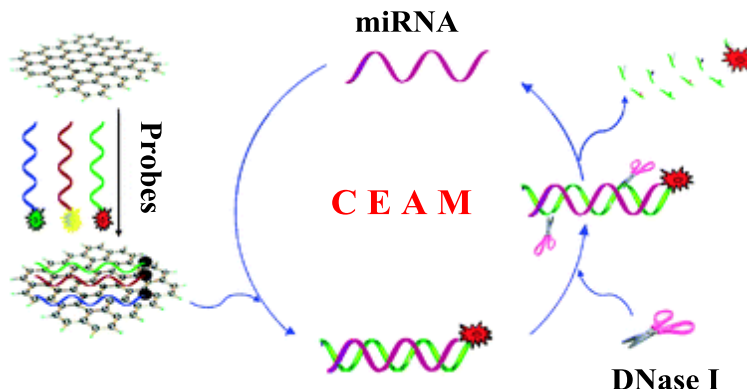


Figure 24 Multiple miRNA detection with signal amplification by cyclic enzyme reaction (adopted from ref^{87d}).

GO-based sensors were developed also to detect protein,⁹⁴ hormone,⁹⁵ adenosine-50- triphosphate (ATP), and fungi toxin⁹⁶ with the use of GO aptamer complexes. The detective role of GO in biomolecules have been extended to the toxic metal ions Hg(II),⁹⁷ Ag (I),⁹⁸ Cu(II),⁹⁹ and Pb(II)^{88a} using GO complexes with fluorescent dye, cystein-rich ssDNA, or combination of ssDNAs and DNAzymes. Recently, the GO sensor has been used for the real-time activity assay of the duplex-unwinding helicase enzyme by monitoring the fluorescence quenching of the ssDNA generated from the helicase unwinding reaction of dsDNA containing a fluorescent dye at the end of one strand.¹⁰⁰ After the development of this system, there has been many attempts to develop enzyme assay systems for DNA dependent protein kinase,¹⁰¹ endonuclease/methyl transferase,¹⁰² exonuclease,¹⁰³ and telomerase,¹⁰⁴ all of which catalyze the biological reaction of NAs. Also, by using fluorescent dye labeled peptides as probes, activities of several important proteases, such as trypsin,¹⁰⁵ thrombin,¹⁰⁶ and metalloproteinase-2,¹⁰⁷ were successfully monitored by protease ss-induced cleavage at specific sites of the peptide probes, resulting in fluorescence recovery. In addition, GO biosensing platforms have been applied for monitoring the Caspase-3 activity¹⁰⁸ and ATP detection¹⁰⁹ in live cells.

GO-organic dye ionic complex was developed to detect ds DNA through ionic exchange on the carboxylic acid groups at the edges of GO. Where GO is complexed with positively charged organic dye, 4-(1-pyrenylvinyl)- N -butylpyridinium bromide (PNPB) and whose fluorescence was quenched upon the formation of PNP + GO – complex. The quenched fluorescence of the dye will be turned on when the GO got exchanged with the DNA from PNP + GO – complex. This exchange was shown to be specific with DNA among the common biomolecules like RNA, glucose, and proteins.¹¹⁰

In summary, the unique and fascinating properties of graphene derivatives such as functionalizable surfaces, strong UV absorption, fluorescence and fluorescence quenching ability make them one of the most promising materials for biosensors, therapeutics, and tissue engineering as well as electronics. In sensing applications, graphene-based materials, featured with good conductivity, large specific surface area and functionalizable surfaces, have demonstrated accurate, rapid, selective, sensitive, and even single-molecular sensing abilities. Although several challenging issues remain, initial promising results in these areas point toward significant potential for graphene derivatives in biomedical research.

1D Origin of DNA/RNA

Many different aspects of chemistry are brought together in this interdisciplinary research such as synthesis of nucleosides and nucleotides, the assembly of these small building blocks into information-containing polymers (DNA and RNA), the origin of homochirality, the formation of protocells etc.

The Miller-Urey experiment conducted by Stanley L. Miller and Harold C. Urey in 1953 is considered as the classical experiment regarding the experimental abiogenesis.¹¹¹ This experiment demonstrated how the organic molecules could have spontaneously formed from inorganic precursors under Oparin-Haldane hypothesis' reducing conditions. The experiment used a highly reduced mixture of gases—methane, ammonia and hydrogen—and water in a sealed glass flask, simulating the atmosphere through electric sparks. The analysis of the resulting pink coloured mixture was shown to contain racemic amino acids.

Apart from the Miller–Urey experiment, there are few notable discoveries in prebiotic organic synthesis that provided the biomolecules sugar and nucleobases. The first one is the formose reaction- alkaline solution of formaldehyde leading to a mixture of higher order sugars, discovered by Aleksandr Butlerov in 1861.¹¹² Almost after a century of the discovery of the formose reaction, Ronald Breslow proposed its mechanism.¹¹³ Another is the demonstration by Joan Oró- adenine synthesis by heating aqueous ammonium cyanide solutions.¹¹⁴ Finally, the HCN oligomerization via the photochemical rearrangement leading to adenine,¹¹⁵ was thought as an important discovery in the ‘RNA world’ hypothesis and also a key point to Eschenmoser’s ‘Glyoxylate scenario’.¹¹⁶ In the ‘glyoxylate scenario’, Eschenmoser proposed “aquo-oligomers of carbon monoxide”, such as glyoxylate and dihydroxyfumarate as central starting materials of a possible primordial metabolism, serving as source molecules for the formation of biogenic molecules such as sugars, α -amino acids, pyrimidines, and constituents of the citric acid cycle. His hypothesis offers an additional pathway to the formation of carbohydrates besides the classical formose reaction with formaldehyde as a source. Recently, Eschenmoser’s hypothesis was strengthened by the experimental observation of the cyanide mediated glyoxylate dimerization to tartarates (via dihydroxyfumarate)¹¹⁷ which could lead to pyruvate, a feedstock molecule to enter into the citric acid cycle. Previously dihydroxyfumarate has been shown to react with small molecule aldehydes leading to selective formation of ketoses (dihydroxy acetone, tetulose and pentuloses).¹¹⁸

In the RNA world hypothesis, assembling the fundamental building blocks (sugar, nucleobases and phosphate) to RNA has proven to be notoriously difficult.¹¹⁹ Another approach would be to hypothesize that life did not begin with RNA; some other genetic system preceded RNA during evolution. For genetic information to be transmitted, the sequence of hydrogen bond donors and acceptors must be ordered. In the pre-RNA world, the alternative nucleic acids which are supposed to precede RNA, should be capable of communicating through the informational Watson-Crick base-pairing. These alternative nucleic acids should also be synthesized by the same system’s chemistry which created RNA, nature’s selection of information polymer. The alternative nucleic acids might have

resulted from the variations in the constitutions of canonical ribose sugar, nucleobases and 5'-3' phosphodiester linkage. Among these many probabilities, the potential natural nucleobase alternatives are much fewer than the possible backbones and sugars.

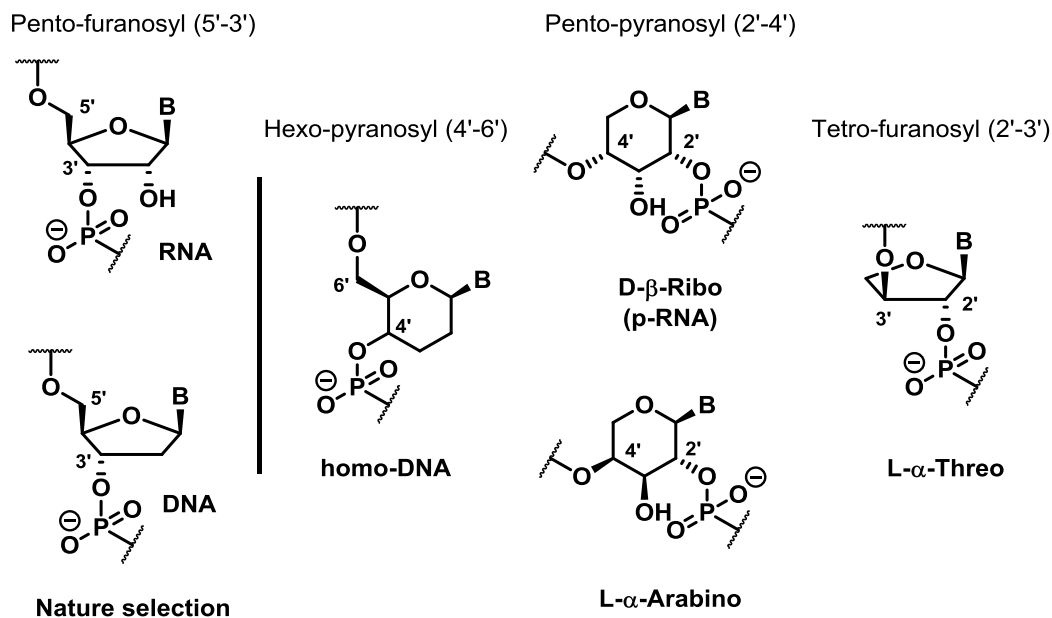


Figure 25 Natural and alternative nucleic acids capable of Watson-Crick base-pairing.

A systematic investigation of potentially natural nucleic acid analogues containing various sugars and linkage isomers has led to the recognition of some intriguing pairing systems (Figure 25).¹²⁰ Eschenmoser's question about nature's selection, 'why pentose and not hexose nucleic acids?' has led to the screening of several sugars (4-6 carbons). The ensuing systematic experimental exploration of (D)-allo-, or altro-, or gluco-hexopyranose's inability in efficient informational base-pairing, unlike the model system 'homo-DNA' (Figure 25) has answered the question 'why pentose and not hexose nucleic acids?'.^{120a} As the corresponding model system 'homo-DNA' has been found to show efficient Watson-Crick pairing, the incapability of hexopyranose's base-pairing has been attributed to the steric bulk (presence of two additional hydroxyl groups).¹²¹ The inference of steric bulk affecting the hexose base-pairing properties narrows down Eschenmoser's

question to the lower order sugars- pentoses and tetroses. Even among pentoses, ‘If ribose, why ribofuranose and not ribopyranose?’ has been asked.

An extensive investigation of structural and chemical properties was extended to the whole family of diastereoisomeric pentopyranosyl- (2'-4') oligonucleotide systems containing D- β -ribo- (p-RNA), D- β -xylo-, L- α -lyxo-, and L- α -arabinopyranosyl units as building blocks, having nucleobases equatorially positioned on the pyranosyl chairs. All the diastereomeric pentopyranosyls studied were stronger Watson-Crick base-pairing systems than hexopyranosyls base-pairing, α -arabinopyranosyl (Figure 25) being the strongest oligonucleotide- type base-pairing system in the family.¹²² (L)- α -threofuranosyl oligonucleotide system (TNA) (Figure 25), the structurally simpler analogue, derived from only four carbons, also exhibited Watson-Crick base-pairing. From these experimental investigations, it is clear that nature did not choose RNA according to the criterion of maximizing base-pairing strength, because the alternative nucleic acids taken from the structural neighborhood of RNA exhibited much stronger pairing properties.

If the predecessor of RNA was able to cross-pair with RNA, the transition may have been a gradual one, and the genetic information could have been preserved by ‘transcription’ of the pre-RNA to RNA. Once the RNA became self-replicating, it took the role of genetic material and the pre-RNA might have been dispensed during evolution. In conclusion, the story of searching for the origins of life (RNA/ pre-RNA), has witnessed several hypotheses, experiments and conjectures and is still persisting.

1E The present work

The following chapters of the thesis illustrate the design and synthesis of different chemically modified DNA/RNA analogues/mimics towards the ultimate goal of therapeutics (splice corrections and miRNA targeting), gene diagnostics in conjunction with graphene oxide and also in the context of origin of life (Pre-RNA world).

In brief it is divided as follows:

Chapter 2 This Chapter has been divided into three sections

Chapter1

Section A: In this section we describe the design, synthesis and properties of 2'-*O*-[*R*- and *S*-(2-amino-3-methoxy)propyl] (*R*-AMP and *S*-AMP) nucleic acids and also an efficient synthetic route for 2'-functionalization of ribopyrimidine nucleosides.

Section B: In this section we describe the synthesis of designed serinol-end-caps for the protection of antisense oligonucleotides from enzymatic degradation. Further, synthesis and evaluation of serinol-capped 2'-*OMe*-mixmers containing T^{*R*-AMP} monomers for splice correction.

Section C: In this section we describe the synthesis of designed 2'-guanidino-functionalized oligonucleotides and evaluation of the biophysical properties.

Chapter 3 This Chapter has been divided into two sections

Section A: In this section we describe the synthesis of designed *S*-polycarbamate nucleic acid and the evaluation of binding properties with complementary DNA/RNA.

Section B: In this section we describe the interactions of different synthetic DNA/RNA analogues with GO and attempts to develop better probe for DNA diagnostics.

Chapter 4 The experimental observation that the quantitative formation of erythrulose from the reaction of glyoxylate and glycoaldehyde-, has led us to design a new acyclic nucleoside analogue based on erythrulose, a four carbon ketose. This chapter describes the design and synthesis of erythrulose based thymine and adenine monomers, and the attempted synthesis of modified oligonucleotides.

1F References

1. Takahashi, I.; Marmur, J., *Nature* **1963**, 197, 794-795.
2. Knowles, J., *Annu. Rev. Biochem.* **1980**, 49, 877-919.
3. Hoogsteen, K., *Acta. Cryst.* **1963**, 65, 907.
4. Crick, F. H. C., *J. Mol. Biol.* **1966**, 19, 548-555.
5. Seeman, N.; Rosenberg, J.; Rich, A., *Proc. Natl. Acad. Sci. USA* **1976**, 73, 804-808.
6. Lescrinier, E.; Froeyen, M.; Herdewijn, P., *Nucleic Acids Res.* **2003**, 31, 2975-2989.
7. Thibaudeau, C.; Acharya, P.; Chattopadhyaya, J., *Stereoelectronic effects in nucleosides and nucleotides and their structural implications*, 2005 (2nd edition), Uppasala University Press. pp 166.
8. (a)Richmond, T. J.; Davey, C. A., *Nature* **2003**, 423, 145-150. (b)Leslie, A. G. W.; Arnott, S.; Chandrasekaran, R.; Ratliff, R. L., *J. Mol. Biol.* **1980**, 143, 49-72. (c)Wahl, M.; Sundaralingam, M., *Biopolymers* **1997**, 44, 45-63. (d)Ghosh, A.; Bansal, M., *Acta Crystallogr D Biol Crystallogr.* **2003**, 59, 620-626.
9. Wang, A. H. J.; Quigley, G. J.; Kalpaks, F. J.; Vander, M. G.; VanBoom, J. H.; Rich, A., *Science* **1981**, 211, 171-176.
10. (a)Hayashi, G.; Hagihara, M.; Nakatani, K., *Nucleic Acids Symp Ser. (Oxf)* **2005**, 49, 261-262. (b)Vargason, J. M.; Eichman, B. F.; Ho, P. S., *Nat. Struct. Biol.* **2000**, 7, 758-761. (c)Wang, G.; Vasquez, K. M., *Mutat. Res.* **2006**, 598, 103-119. (d)Allemand, J. F.; Bensimon, D.; Lavery, R.; Croquette, V., *Proc. Natl. Acad. Sci. USA* **1998**, 95, 14152-14157.
11. Berg, J. M.; Tymoczko, J. L.; Stryer, L., *Biochemistry*, 5th Edition, WH Freeman and Company. **2002**, 781-808.
12. Eschenmoser, A.; Dobler, M., *Helv. Chim. Acta* **1992**, 75, 218-259.
13. Felsenfeld, G.; Davies, D.; Rich, A., *J. Am. Chem. Soc.* **1957**, 79, 2023-2024.
14. (a)Moser, H.; Dervan, P., *Science* **1987**, 238, 645-650. (b)LeDoan, T.; Perrouault, L.; Praseuth, D.; Habhoub, N.; Decout, J.-L.; Thuong, N. T.; Lhomme, J.; He´ne, C., *Nucleic Acids Res.* **1987**, 15, 7749-7760.

15. (a)Thuong, N. T.; Helene, C., *Angew. Chem. Int. Ed.* **1993**, 32, 666-690. (b)Beal, P. A.; Dervan, P. B., *Science* **1991**, 251, 1360-1363.
16. Hoogsteen, K., *Acta. Cryst.* **1959**, 12, 822-823.
17. Limongelli, V.; De Tito, S.; Cerofolini, L.; Fragai, M.; Pagano, B.; Trotta, R.; Cosconati, S.; Marinelli, L.; Novellino, E.; Bertini, I.; Randazzo, A.; Luchinat, C.; Parrinello, M., *Angew. Chem. Int. Ed.* **2013**, 52, 2269-2273.
18. (a)Wells, R. D.; Harvey, S. C., Springer-Verlag, New York. **1988**. (b)Helene, C., *Anti-Cancer Drug Design* **1991**, 6, 569-584. (c)Frank-Kamenetskii, M. D., *Methods in Enzymology* **1992**, 211, 180-191. (d)Soyfer, V. N.; Potaman, V. N., *Eds* **1996**, Springer-Verlag, New York.
19. (a)Bončina, M.; Lah, J.; Prislán, I.; Vesnaver, G., *J. Am. Chem. Soc.* **2012**, 134, 9657-9663. (b)Gray, R.; Buscaglia, R.; Chaires, J., *J. Am. Chem. Soc.* **2012**, 134, 16834-16844. (c)Mashimo, T.; Yagi, H.; Sannohe, Y.; Rajendran, A.; Sugiyama, H., *J. Am. Chem. Soc.* **2010**, 132, 14910-14918.
20. Greider, C. W., *Proc. Natl Acad. Sci. USA* **1998**, 95, 90-92.
21. Makarov, V. L.; Hirose, Y.; Langmore, J. P., *Cell* **1997**, 88, 657-666.
22. (a)Sundquist, W. I.; Klug, A., *Nature* **1989**, 342, 825-829. (b)Sen, D.; Gilbert, W., *Nature* **1988**, 334, 364-366.
23. Gehring, K.; Leroy, J. L.; Guéron, M., *Nature* **1993**, 363, 561-565.
24. (a)Burge, S.; Parkinson, G.; Hazel, P.; Todd, A.; Neidle, S., *Nucleic Acids Res.* **2006**, 34, 5402-5415. (b)Bates, P.; Mergny, J.-L.; Yang, D., *EMBO Reports* **2007**, 8, 1003-1010. (c)Collie, G.; Parkinson, G., *Chem. Soc. Rev.* **2011**, 40, 5867-5892.
25. Blackburn, G. M.; Gait, M. J.; Loakes, D.; Williams, D. M., *Nucleic acids in chemistry and biology*, RSC publishing, 3rd edition.
26. Cantor, C. R.; Schimmel, P. R., *Biophysical Chemistry part III*, **1971**, W. H. Freeman and Company, New York.
27. Plum, G. E.; Park, Y.-H.; Singleton, S. F.; Dervan, P. B.; Breslauer, K. J., *Proc. Nat. Acad. Sci. USA* **1990**, 87, 9436-9440.
28. Van Holde, K. E.; Brahm, J.; Michelson, M. E., *J. Mol. Biol.* **1965**, 12, 726-739.

29. (a)Egli, M.; Williams, L. D.; Gao, Q.; Rich, A., *Biochemistry* **1991**, 30, 11388-11402. (b)Gassner, R. V.; Frederick, C. A.; Quigley, G. J.; Rich, A.; Wang, A. H., *J. Biol. Chem.* **1989**, 264, 7921-7935.
30. Zamecnik, P. C.; Stephenson, M. L., *Proc. Natl. Acad. Sci. USA* **1978**, 75, 280-284.
31. Agrawal, S.; Zhao, Q. Y., *Curr. Opin. Chem. Biol.* **1998**, 2, 519-528.
32. Raal, F. J.; Santos, R. D.; Blom, D. J.; Marais, A. D.; Charng, M.-J.; Cromwell, W. C.; Lachmann, R. H.; Gaudet, D.; Tan, J. L.; Chasan-Taber, S.; Tribble, D. L.; Flaim, J. D.; Crooke, S. T., *The Lancet* **2010**, 375, 998-1006.
33. Davidson, B. L.; McCray Jr, P. B., *Nat. Rev. Genet.* **2011**, 12, 329-340.
34. Smith, C. C.; Aurelian, L.; Reddy, M. P.; Miller, P. S.; Ts'o, P. O., *Proc. Natl Acad. Sci. USA* **1986**, 83, 2787-2791.
35. Stirchak, E. P.; Summerton, J. E.; Weller, D. D., *Nucleic Acids Res.* **1989**, 17, 6129-6141.
36. Fire, A.; Xu, S.-Q.; Montgomery, M. K.; Kostas, S. A.; Driver, S. E.; Mello, C. C., *Nature* **1998**, 391, 806-811.
37. (a)Aravin, A.; Gaidatzis, D.; Pfeffer, S.; Lagos-Quintana, M.; Landgraf, P.; Iovino, N.; Morris, P.; Brownstein, M. J.; Kuramochi-Miyagawa, S.; Nakano, T.; Chien, M.; Russo, J. J.; Ju, J.; Sheridan, R.; Sander, C.; Zavolan, M.; Tuschl, T., *Nature* **2006**, 442, 203-207. (b)Girard, A.; Sachidanandam, R.; Hannon, G. J.; Carmell, M. A., *Nature* **2006**, 442, 199-202. (c)Grivna, S. T.; Beyret, E.; Wang, Z.; Lin, H., *Genes & Development* **2006**, 20, 1709-1714. (d)Watanabe, T.; Takeda, A.; Tsukiyama, T.; Mise, K.; Okuno, T.; Sasaki, H.; Minami, N.; Imai, H., *Genes & Development* **2006**, 20, 1732-1743.
38. Elbashir, S. M.; Harborth, J.; Lendeckel, W.; Yalcin, A.; Weber, K.; Tuschl, T., *Nature* **2001**, 411, 494-498.
39. Matranga, C.; Tomari, Y.; Shin, C.; Bartel, D. P.; Zamore, P. D., *Cell* **2005**, 123, 607-620.
40. Fougerolles, A.-d.; Vornlocher, H.-P.; Maraganore, J.; Lieberman, J., *Nat. Rev. Drug Discovery* **2007**, 6, 443-453.

41. Liu, J. D.; Carmell, M. A.; Rivas, F. V.; Marsden, C. G.; Thomson, J. M.; Song, J. J.; Hammond, S. M.; Joshua-Tor, L.; Hannon, G. J., *Science* **2004**, 305, 1437-1441.
42. Song, J. J.; Smith, S. K.; Hannon, G. J.; Joshua-Tor, L., *Science* **2004**, 305, 1434-1437.
43. Lee, R. C.; Feinbaum, R. L.; Ambros, V., *Cell* **1993**, 75, 843-854.
44. Kim, V. N., *Nat. Rev. Mol. Cell Biol.* **2005**, 6, 376-385.
45. Garofalo, M.; Croce, C. M., *Annu. Rev. Pharmacol. Toxicol.* **2011**, 51, 25-43.
46. Molecular Biology Select. *Cell* **2006**, 126, 223-225.
47. Gilbert, W., *Nature* **1978**, 271, 501-501.
48. Matlin, A. J.; Clark, F.; Smith, C. W. J., *Nat. Rev. Mol. Cell Biol.* **2005**, 6, 386-398.
49. Black, D. L., *Annu. Rev. Biochem.* **2003**, 72, 291-336.
50. Ozsolak, F.; Milos, P. M., *Nat. Rev. Genet.* **2011**, 12, 87-98.
51. (a)Sazani, P.; Graziewicz, M.; Kole, R., *AntisenseDrug Technology: Principles, Strategies, and Applications* 2nd edn (ed. Crooke, S. T). *CRC Press, Boca Raton* **2007**, 89–114. (b)Sazani, P.; Gemignani, F.; Kang, S.-H.; Maier, M. A.; Manoharan, M.; Persmark, M.; Bortner, D.; Kole, R., *Nat. Biotech.* **2002**, 20, 1228-1233. (c)Roberts, J.; Palma, E.; Sazani, P.; arum, H.; Cho, M.; Kole, R., *Mol. Ther.* **2006**, 14, 471-475.
52. (a)Deutekom, J. C. v.; Janson, A. A.; Ginjaar, I. B.; Frankhuizen, W. S.; Aartsma-Rus, A.; Bremmer-Bout, M.; Dunnen, J. T. d.; Klaas Koop; Kooi, A. J. v. d.; Goemans, N. M.; Kimpe, S. J. d.; Ekhart, P. F.; Venneker, E. H.; Platenburg, G. J.; Verschuuren, J. J.; Ommen, G.-J. B. v., *N. Engl. J. Med.* **2007**, 357, 2677-2686. (b)Kinali, M.; Arechavala-Gomez, V.; Feng, L.; Cirak, S.; Hunt, D.; Adkin, C.; Guglieri, M.; Ashton, E.; Abbs, S.; Nihoyannopoulos, P.; Garralda, M. E.; Rutherford, M.; Mcculley, C.; Popplewell, L.; Graham, I. R.; Dickson, G.; Wood, M. J.; Wells, D. J.; Wilton, S. D.; Kole, R.; Straub, V.; Bushby, K.; Sewry, C.; Morgan, J. E.; Muntoni, F., *Lancet Neurol.* **2009**, 8, 918-928. (c)Cirak, S.; Arechavala-Gomez, V.; Guglieri, M.; Feng, L.; Torelli, S.; Anthony, K.; Abbs, S.; Garralda, M. E.; Bourke, J.; Wells, D. J.; Dickson, G.; Wood, M. J.; Wilton, S. D.;

- Straub, V.; Kole, R.; Shrewsbury, S. B.; Sewry, C.; Morgan, J. E.; Bushby, K.; Muntoni, F., *The Lancet* **2011**, 378, 595-605. (d)Goemans, N. M.; Tulinius, M.; Akker, J. T. v. d.; Burm, B. E.; Ekhart, P. F.; Heuvelmans, N.; Holling, T.; Janson, A. A.; Platenburg, G. J.; Sipkens, J. A.; Sitsen, J. M. A.; Aartsma-Rus, A.; Ommen, G.-J. B. v.; Buyse, G.; Darin, N.; Verschuuren, J.; Campion, G. V.; Kimpe, S. J. d.; Deutekom, J. C. v., *N. Engl. J. Med.* **2011**, 364, 1513-1522.
53. Kang, S. H.; Cho, M. J.; Kole, R., *Biochemistry* **1998**, 37, 6235-6239.
54. Eckstein, F., *Antisense Nucleic Acid Drug Dev.* **2000**, 10, 117-121.
55. Summerton, J., *Biochim. Biophys. Acta* **1999**, 1489, 141-158.
56. Draper, B. W.; Morcos, P. A.; Kimmel, C. B., *Genesis* **2001**, 30, 154-156.
57. (a)Levin, A. A., *Biochim. Biophys. Acta* **1999**, 1489, 69-84. (b)Crooke, S. T., 'Phosphorothioate Oligonucleotides', in: 'Therapeutic Applications of Oligonucleotides', Ed. S. T. Crooke, R. G. Landes, Austin. **1995**, p. 79.
58. Guschlbauer, W.; Jankowski, K., *Nucleic Acids Res.* **1980**, 8, 1421-1433.
59. Viazovkina, E.; Mangos, M. M.; Elzagheid, M. I.; Damha, M. J., *Curr. Protoc. Nucleic. Acid Chem.* Chapter 4, unit 415. **2002**.
60. Watts, J. K.; Martin-Pintado, N.; Gomez-Pinto, I.; Schwartzentruber, J.; Portella, G.; Orozco, M.; Gonzalez, C.; Damha, M. J., *Nucleic Acids Res.* **2010**, 38, 2498-2511.
61. Inoue, H.; Hayase, Y.; Imura, A.; Iwai, S.; Miura, K.; Ohtsuka, E., *Nucleic Acids Res.* **1987**, 15, 6131-6148.
62. Martin, P., *Helv. Chim. Acta* **1995**, 78, 486-504.
63. Prakash, T. P., *Chemistry & biodiversity* **2011**, 8, 1616-1641.
64. Prakash, T. P.; Kawasaki, A. M.; Lesnik, E. A.; Owens, S. R.; Manoharan, M., *Org. Lett.* **2003**, 5, 403-406.
65. Prakash, T. P.; Kawasaki, A. M.; Wancewicz, E. V.; Shen, L.; Monia, B. P.; Ross, B. S.; Bhat, B.; Manoharan, M., *J. Med. Chem.* **2008**, 51, 2766-2776.
66. (a)Singh, S. K.; Nielsen, P.; Koshkin, A. A.; Wengel, J., *Chem. Commun.* **1998**, 455-456. (b)Koshkin, A. A.; Singh, S. K.; Nielsen, P.; Rajwanshi, V. K.; Kumar, R.; Meldgaard, M.; Olsen, C. E.; Wengel, J., *Tetrahedron* **1998**, 54, 3607-3630.

- (c)Obika, S.; Nanbu, D.; Hari, Y.; Morio, K.-i.; In, Y.; Ishida, T.; Imanishi, T., *Tetrahedron Lett.* **1997**, 38, 8735-8738.
67. (a)Koshkin, A. A.; Nielsen, P.; Meldgaard, M.; Rajwanshi, V. K.; Singh, S. K.; Wengel, J., *J. Am. Chem. Soc.* **1998**, 120, 13252-13253. (b)Obika, S.; Nanbu, D.; Hari, Y.; Andoh, J.-i.; Morio, K.-i.; Doi, T.; Imanishi, T., *Tetrahedron Lett.* **1998**, 39, 5401-5404. (c)Wengel, J., *Acc. Chem. Res.* **1999**, 32, 301-310.
68. Singh, S. K.; Kumar, R.; Wengel, J., *J. Org. Chem.* **1998**, 63, 6078-6079.
69. Zhou, C.; Liu, Y.; Andaloussi, M.; Badgujar, N.; Plashkevych, O.; Chattopadhyaya, J., *J. Org. Chem.* **2009**, 74, 118-134.
70. Morita, K.; Takagi, M.; Hasegawa, C.; Kaneko, M.; Tsutsumi, S.; Sone, J.; Ishikawa, T.; Imanishi, T.; Koizumi, M., *Bioorganic & Medicinal Chemistry* **2003**, 11, 2211-2226.
71. Koizumi, M.; Morita, K.; Daigo, M.; Tsutsumi, S.; Abe, K.; Obika, S.; Imanishi, T., *Nucleic Acids Res.* **2003**, 31, 3267-3273.
72. Varghese, O. P.; Barman, J.; Pathmasiri, W.; Plashkevych, O.; Honcharenko, D.; Chattopadhyaya, J., *J. Am. Chem. Soc.* **2006**, 128, 15173-15187.
73. Seth, P. P.; Siwkowski, A.; Allerson, C. R.; Vasquez, G.; Lee, S.; Prakash, T. P.; Kinberger, G.; Migawa, M. T.; Gaus, H.; Bhat, B.; Swayze, E. E., *Nucleic Acids Symp Ser.* **2008**, 52, 553-554.
74. Nielsen, P. E.; Egholm, M.; Berg, R. H.; Buchardt, O., *Science* **1991**, 254, 1497-1500.
75. Egholm, M.; Buchardt, O.; Christensen, L.; Behrens, C.; Freier, S. M.; Driver, D. A.; Berg, R. H.; Kim, S. K.; Nordon, B.; Nielsen, P. E., *Nature* **1993**, 365, 566-568.
76. (a)Ganesh, K. N.; Nielsen, P. E., *Curr. Org. Chem.* **2000**, 4, 931-943. (b)Kumar, V. A., *Eur. J. Org. Chem.* **2002**, 2002, 2021-2032. (c)Kumar, V. A.; Ganesh, K. N., *Acc. Chem. Res.* **2005**, 38, 404-412. (d)Uhlmann, E.; Breipohl, G.; Will, D., *Angew. Chem., Int. Ed.* **1998**, 37, 2796-2823.
77. Zhou, P.; Wang, M.; Du, L.; Fisher, G. W.; Waggoner, A.; Ly, D. H., *J. Am. Chem. Soc.* **2003**, 125, 6878-6879.

78. Gokhale, S. S.; Kumar, V. A., *Org. Biomol. Chem.* **2010**, 8, 3742-3750.
79. (a)Nielsen, P. E.; Egholm, M., *Peptide Nucleic Acids. Protocols and Applications* **1999**, Horizon Press: Norfolk. (b)Ray, A.; Norden, B., *FASEB Journal* **2000**, 14, 1041-1060.
80. Good, L.; Nielsen, P. E., *Nature Biotechnology* **1998**, 16, 355-358.
81. Nielsen, P. E., *Curr. Opin. Biotechnol.* **1999**, 10, 71-75.
82. Metaferia, B.; Wei, J. S.; Song, Y. K.; Evangelista, J.; Aschenbach, K.; Johansson, P.; Wen, X.; Chen, Q.; Lee, A.; Hempel, H.; Gheeya, J. S.; Getty, S.; Gomez, R.; Khan, J., *PLoS ONE* **2013**, 8, e58870.
83. (a)Novoselov, K. S.; Geim, A. K.; Morozov, S. V.; Jiang, D.; Zhang, Y.; Dubonos, S. V.; Grigorieva, I. V.; Firsov, A. A., *Science* **2004**, 306, 666-669. (b)Geim, A. K.; Novoselov, K. S., *Nat. Mater.* **2007**, 6, 183-191.
84. Chung, C.; Kim, Y.-K.; Shin, D.; Ryoo, S.-R.; Hong, B.; Min, D.-H., *Acc. Chem. Res.* **2013**, DOI: 10.1021/ar300159f.
85. Xu, Y.; Bai, H.; Lu, G.; Li, C.; Shi, G., *J. Am. Chem. Soc.* **2008**, 130, 5856-5857.
86. (a)Swathi, R. S.; Sebastian, K. L., *J. Chem. Phys.* **2009**, 130, 86101. (b)Varghese, N.; Mogera, U.; Govindaraj, A.; Das, A.; Maiti, P.; Sood, A.; Rao, C. N. R., *Chem. Phys. Chem.* **2009**, 10, 206-210. (c)Liu, Z.; Robinson, J. T.; Sun, X.; Dai, H., *J. Am. Chem. Soc.* **2008**, 130, 10876-10877.
87. (a)Lu, C.-H.; Yang, H.-H.; Zhu, C.-L.; Chen, X.; Chen, G.-N., *Angew. Chem. Int. Ed.* **2009**, 48, 4785-4787. (b)Wu, C.; Zhou, Y.; Miao, X.; Ling, L., *The Analyst* **2011**, 136, 2106-2110. (c)Lu, Z.; Zhang, L.; Deng, Y.; Li, S.; He, N., *Nanoscale* **2012**, 4, 5840-5842. (d)Cui, L.; Lin, X.; Lin, N.; Song, Y.; Zhu, Z.; Chen, X.; Yang, C. J., *Chem. Commun.* **2012**, 48, 194-196.
88. (a)Zhao, X.-H.; Kong, R.-M.; Zhang, X.-B.; Meng, H.-M.; Liu, W.-N.; Tan, W.; Shen, G.-L.; Yu, R.-Q., *Anal. Chem.* **2011**, 83, 5062-5066. (b)Huang, W. T.; Xie, W. Y.; Shi, Y.; Luo, H. Q.; Li, N. B., *J. Mater. Chem.* **2012**, 22, 1477-1481. (c)Xie, W. Y.; Huang, W. T.; Li, N. B.; Luo, H. Q., *Chem. Commun.* **2012**, 48, 82-84.

89. (a)Xu, F.; Shi, H.; He, X.; Wang, K.; Ye, X.; Yan, L. a.; Wei, S., *Analyst* **2012**, 137, 3989-3994. (b)Zhou, Z.; Zhu, C.; Ren, J.; Dong, S., *Anal. Chim. Acta* **2012**, 740, 88-92.
90. Li, F.; Feng, Y.; Zhao, C.; Li, P.; Tang, B., *Chem. Commun.* **2012**, 48, 127-129.
91. Shao, Y.; Zhang, S.; Engelhard, M. H.; Li, G.; Shao, G.; Wang, Y.; Liu, J.; Aksay, I. A.; Lin, Y., *J. Mater. Chem.* **2010**, 20, 7491-7496.
92. Huang, X.; Yin, Z.; Wu, S.; Qi, X.; He, Q.; Zhang, Q.; Yan, Q.; Boey, F.; Zhang, H., *Small* **2011**, 7, 1876-1902.
93. He, S.; Song, B.; Li, D.; Zhu, C.; Qi, W.; Wen, Y.; Wang, L.; Song, S.; Fang, H.; Fan, C. A., *Adv. Funct. Mater.* **2010**, 20, 453-459.
94. Chang, H.; Tang, L.; Wang, Y.; Jiang, J.; Li, J., *Anal. Chem.* **2010**, 82, 2341-2346.
95. Pu, Y.; Zhu, Z.; Han, D.; Liu, H.; Liu, J.; Liao, J.; Zhang, K.; Tan, W., *Analyst* **2011**, 136, 4138-4140.
96. Sheng, L.; Ren, J.; Miao, Y.; Wang, J.; Wang, E., *Biosens. Bioelectron.* **2011**, 26, 3494-3499.
97. Huang, W. T.; Shi, Y.; Xie, W. Y.; Luo, H. Q.; Li, N. B., *Chem. Commun.* **2011**, 47, 7800-7802.
98. Wen, Y.; Xing, F.; He, S.; Song, S.; Wang, L.; Long, Y.; Li, D.; Fan, C., *Chem. Commun.* **2010**, 46, 2596-2598.
99. Liu, M.; Zhao, H.; Chen, S.; Yu, H.; Zhang, Y.; Quan, X., *Chem. Commun.* **2011**, 47, 7749-7751.
100. Jang, H.; Kim, Y.-K.; Kwon, H.-M.; Yeo, W.-S.; Kim, D.-E.; Min, D.-H., *Angew. Chem. Int. Ed.* **2010**, 49, 5703-5707.
101. Lin, L.; Liu, Y.; Zhao, X.; Li, J., *Anal. Chem.* **2011**, 83, 8396-8402.
102. Lee, J.; Kim, Y.-K.; Min, D.-H., *Anal. Chem.* **2011**, 83, 8906-8912.
103. Lee, J.; Min, D.-H., *Analyst* **2012**, 137, 2024-2026.
104. Peng, L.; Zhu, Z.; Chen, Y.; Han, D.; Tan, W., *Biosens. Bioelectron.* **2012**, 35, 475-478.

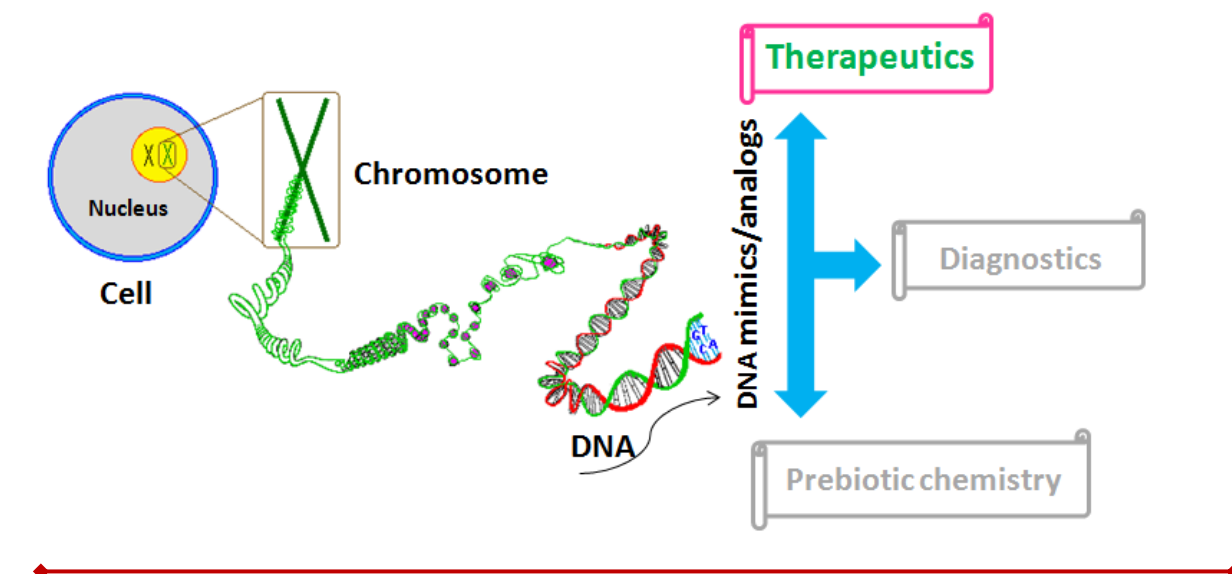
105. Gu, X.; Yang, G.; Zhang, G.; Zhang, D.; Zhu, D., *ACS Appl. Mater. Interfaces*. **2011**, 3, 1175-1179.
106. Zhang, M.; Yin, B.-C.; Wang, X.-F.; Ye, B.-C., *Chem. Commun.* **2011**, 47, 2399-2401.
107. Feng, D.; Zhang, Y.; Feng, T.; Shi, W.; Li, X.; Ma, H., *Chem. Commun.* **2011**, 47, 10680-10682.
108. Wang, H.; Zhang, Q.; Chu, X.; Chen, T.; Ge, J.; Yu, R., *Angew. Chem. Int. Ed.* **2011**, 50, 7065-7069.
109. Wang, Y.; Li, Z.; Hu, D.; Lin, C.-T.; Li, J.; Lin, Y., *J. Am. Chem. Soc.* **2010**, 132, 9274-9276.
110. Balapanuru, J.; Yang, J.-X.; Xiao, S.; Bao, Q.; Jahan, M.; Polavarapu, L.; Wei, J.; Xu, Q.-H.; Loh, K. P., *Angew. Chem. Int. Ed.* **2010**, 49, 6549-6553.
111. (a)Miller, S. L., *Science* **1953**, 117, 528-529. (b)Miller, S. L.; Urey, H. C., *Science* **1959**, 130, 245-251.
112. Boutlerow, A., *Justus Liebigs Annalen der Chemie* **1861**, 120, 295-298.
113. Breslow, R., *Tetrahedron Letters* **1959**, 1, 22-26.
114. Oró, J., *Nature* **1961**, 191, 1193-1194.
115. Ferris, J. P.; Orgel, L. E., *J. Am. Chem. Soc.* **1966**, 88, 1074.
116. (a)Eschenmoser, A., *Tetrahedron* **2007**, 63, 12821-12844. (b)Eschenmoser, A., *Chemistry & Biodiversity* **2007**, 4, 554-573.
117. Butch, C.; Cope, E. D.; Pollet, P.; Gelbaum, L.; Krishnamurthy, R.; Liotta, C. L., *J. Am. Chem. Soc.* **2013**, 135, 13440-13445.
118. Sagi, V. N.; Punna, V.; Hu, F.; Meher, G.; Krishnamurthy, R., *J. Am. Chem. Soc.* **2012**, 134, 3577-3589.
119. Orgel, L. E., *Crit. Rev. Biochem. Mol. Biol.* **2004**, 39, 99-123.
120. (a)Eschenmoser, A., *Science* **1999**, 284, 2118-2124. (b)Eschenmoser, A., *Angew. Chem. Int. Ed.* **2011**, 50, 12412-12472. (c)Engelhart, A. E.; Hud, N. V., *Cold Spring Harb. Perspect. Biol.* **2010**, 2:a002196. (d)Joyce, G. F., *Nature* **2002**, 418, 214-221. (e)Loakes, D.; Holliger, P., *Molecular BioSystems* **2009**, 5, 686-694.

Chapter1

121. Egli, M.; Pallan, P. S.; Pattanayek, R.; Wilds, C. J.; Lubini, P.; Minasov, G.; Dobler, M.; Leumann, C. J.; Eschenmoser, A., *J. Am. Chem. Soc.* **2006**, 128, 10847-10856.
122. Beier, M.; Reck, F.; Wagner, T.; Krishnamurthy, R.; Eschenmoser, A., *Science* **1999**, 283, 699-703.

CHAPTER 2

Applications of synthetic DNA in therapeutics



Section 2A: Design, synthesis and biophysical evaluation of 2'-O-(R/S-serinol) functionalized oligonucleotides

2A.1 Introduction

Synthetic antisense or siRNA oligonucleotides (ONs) as gene silencing agents inhibit viral replication and expression of disease-causing genes based on the simple concept of nucleic acid sequence recognition by a complementary base sequence *via* Watson-Crick hydrogen bonding.¹ In RNA therapeutics, most of the designed synthetic antisense oligonucleotides' primary objective is to increase the duplex stability with the target RNA. The basic chemical difference between DNA and RNA is the absence or presence of 2'-OH substitution respectively, which allows the sugars to adopt different conformations- C2'-*endo* (S-type with extended backbone) in DNA and C3'-*endo* (N-type with compact backbone) in RNA (Figure 1). The observation from the natural nucleic acids that the pseudoaxially oriented nucleobases in the C3'-*endo* conformation of RNA improved the duplex stability with effective base-pair stacking, was the key to the development of 2'- modified sugar in oligonucleotides. Thus, the structural design of many sugar-modified AONs involve the 2'-derivatization by electronegative substituents/2'-O-functionalization to retain the N-type sugar conformation, or to lock the sugar in required N-type conformation by a chemical bridge (Figure 1).

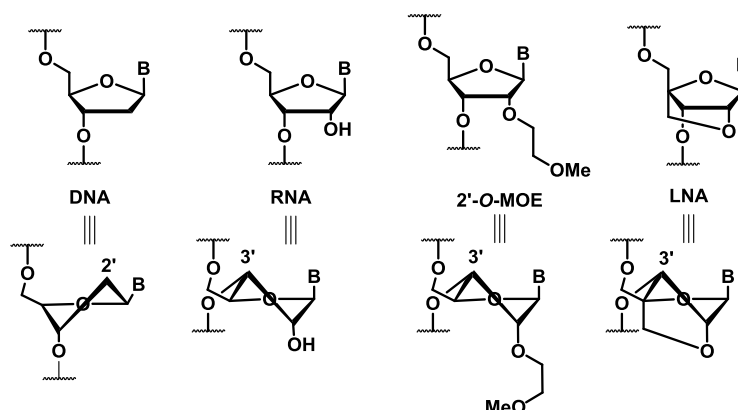


Figure 1 Nucleic acids and analogues in the increasing order of N-type sugar conformation and so the duplex stability with the target RNA (from left to right).

Chapter 2

Amongst the large number of 2'-*O*-alkyl derivatives studied so far (Section 1B.7, Chapter 1),² the 2'-*O*-methoxyethyl substituted antisense oligonucleotides (MOE-AONs) are being studied³ in several ongoing clinical trials and have shown excellent safety profiles. The enhanced duplex stability of MOE AONs with RNA was attributed to the preferred C3'-*endo* conformation (due to the *gauche* orientation of ethylene side chain, confirmed by the X-ray studies⁴) and also conformational pre-organization and extensive hydration in the minor groove. Another promising modified AON is the locked nucleic acid (LNA) and analogues, which imparts excellent duplex stability upon binding to the target RNA (though leading sometimes to off-target effects⁵) and enhanced enzymatic stability. LNA technology has been extensively studied for successful applications in different areas of research such as antisense and antigene strategies, diagnostics, genotyping, and has also progressed to clinical trials.⁶

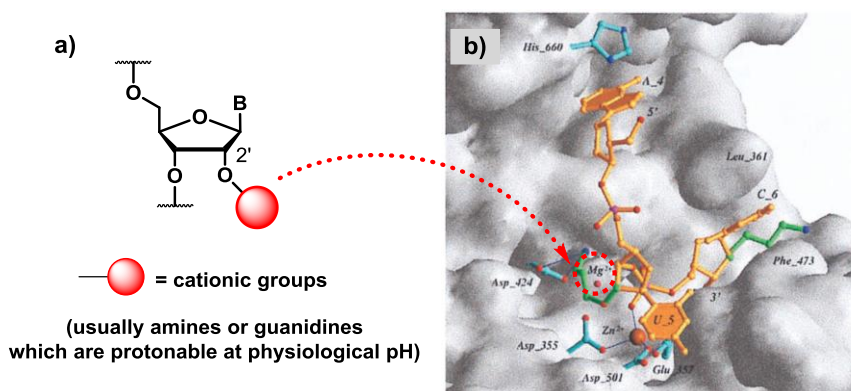


Figure 2 Chemical structure of representative 2'-*O*-tethered cationic group in RNA (a) and the structural model suggesting the interference of the positively charged 2'-*O*-aminopropyl-substituent with the metal ion binding site B of the exonuclease wt-KF (adopted from ref⁷).

It is known that the presence of amino functionality in the minor groove (Figure 2a) would reduce electrostatic charge-repulsions during duplex formation⁸ and also prove useful in reducing the overall negative charges of the ONs that would in turn, improve the cellular uptake of the oligonucleotides.⁹ The 2'-*O*-aminoethyl/2'-*O*-aminopropyl derived DNA oligomers form less stable duplexes with complementary RNA, compared to -OMe or MOE-AONs,^{2b, 10} but show increased stability towards hydrolytic enzymes.¹⁰

Chapter 2

Crystallographic study of the 2'-*O*-aminopropyl- modified AONs bound at the active site of an exonuclease revealed that the protonated amino group of this modification displaced a catalytically important metal cation of the enzyme (Figure 2b).⁷ The modeling studies and crystal structure analysis indicate that the amino group in 2'-*O*-aminopropyl substituted ssDNA does not have effective interactions with the neighbouring 3'-phosphate oxygen, the neighbouring base carbonyl or sugar *O4'* atoms but does provide water molecules spanning the minor groove and improves hydration.^{7, 10} In addition, amino pendant groups have advantages in improved kinetics of binding.¹¹

2A.2 Our design and rationale

Out of the large pool of 2'-alkyl substituted oligomers designed so far,^{2a, 2b, 10} the MOE substitution imparts a high degree of conformational pre-organization through a stable water structure as observed in the crystal structure.^{4a} We designed a novel modification characterized by the presence of a positively charged functionality in conjunction with the methoxy substituent of the 2'-*O*-MOE chain to ensure that the resulting oligomers have advantages of both MOE- and AP-groups (Figure 3). We chose to synthesize derivative A in this work as a part of serinol based project.

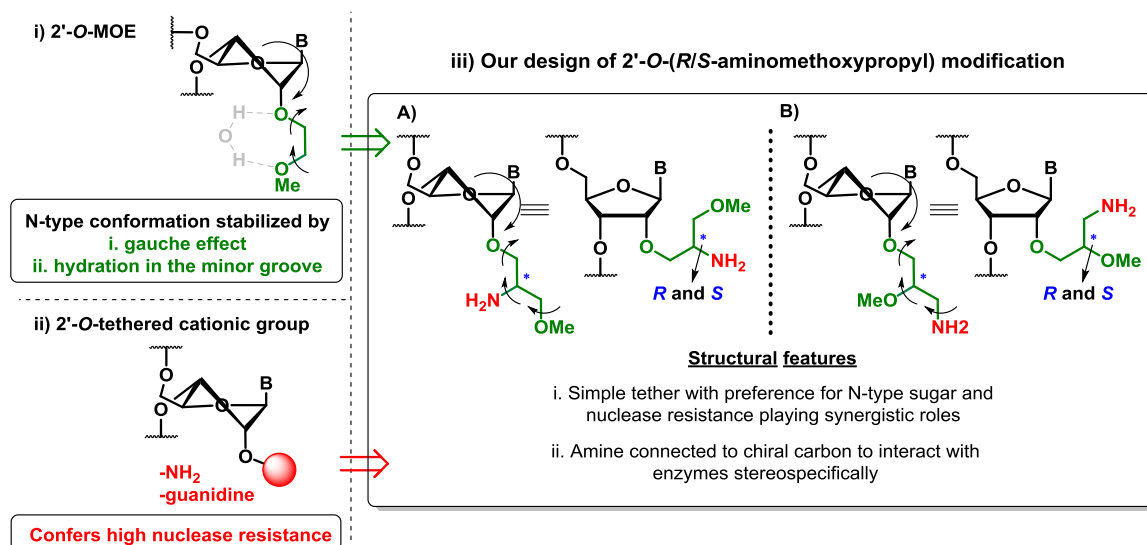


Figure 3 Design of 2'-*O*-(*R/S*-aminomethoxypropyl) modification (iii) featuring the properties of 2'-*O*-MOE (i) and 2'-*O*-tethered cationic groups (ii).

The aminomethoxypropyloxy (AMP) substitution at 2'-position derived from the combination of MOE and AE/AP DNA analogues. Although a large number of 2'-*O*-chemical modifications are reported in the literature, this is perhaps the first time where a stereo-pure branched chain is conjugated to the 2'-*O*- position of AONs. In addition to the presence of charged amino group, the hydrolytic susceptibility of the functionalized AONs is also known to be reduced with increasing size of the 2'-*O*-substitution.^{2b} For the present enantiomeric *R/S*-AMP tethers, due to the conformational freedom and no anticipated interactions with the adjoining phosphate group, carbonyl or O4' oxygen,^{7, 10} the steric bulk and electronic factors may be expected to be similar for both the isomers when incorporated in the ssDNA.¹² Thus, the introduction of an aminoethyl functionality as well as retaining the methoxy group in the 2'-pendant might provide synergy of effects *i. e.* shielding of phosphate group by the positively charged moiety as well as duplex stability resulting from a stable solvent network. We undertook the synthesis of two diastereomeric nucleoside analogues differing in chirality at the 2'-*O*-substituent groups and characterized by the synergistic combination of MOE and amino pendant groups.

2A.3 An efficient synthesis of 2'-*O*-functionalized pyrimidine nucleosides

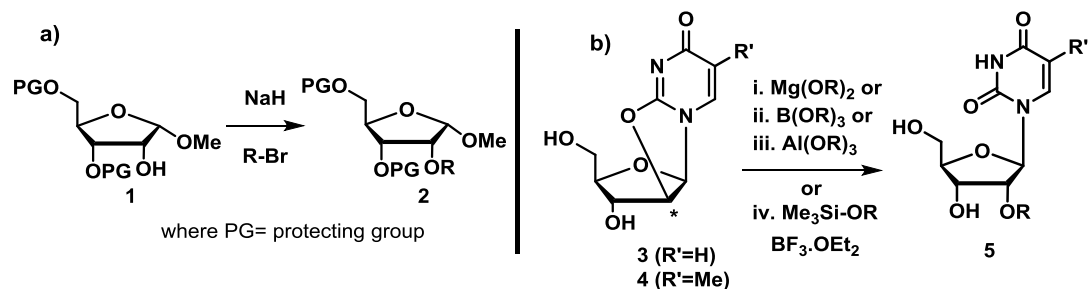


Figure 4 Common approaches for the 2'-*O*-alkylation either by extensive usage of protecting groups (a) or through double activation (electrophile and nucleophile) (b).

There are previous literature reports^{2a, 2b} to achieve the selective alkylation of uridine at 2'-position such as treatment of uridine with alkyl halides using DBTO (dibutyltin oxide)¹³ or by employing 3',5'-*O*- and *N*-protected uridine in reactions with alkyl halides^{2a} or cyanoethylation *via* a Michael addition of 3',5'-*O*-protected uridine to acrylonitrile.¹⁴ Another approach was to start from the sugars and convert them to the 2'-*O*-

functionalized nucleosides *via* extensive protection/deprotection strategy en route (Figure 4a).^{10, 15} The limitation of this approach was the long synthetic route, which dictate the overall yield.

Alternatively, regiospecific 2,2'-anhydro uridine (**3**) ring opening was effected by reaction with borate esters generated from BH_3 ¹⁶ and large excess of appropriate alcohols to get the desired products at high temperature and pressure or by treatment with metal alkoxides (Figure 4b).¹⁷ The yields were variable in each case and in the case of *N*-2-hydroxyethylphthalimide, the reported yield was 21%.¹⁸ The main limiting factor of this methodology in the present work is 2'-*O*-branched functionalization, where large excess alcohols need to be synthesized (almost like a solvent). Alternatively, we came across $\text{BF}_3 \cdot \text{OEt}_2$ - promoted 2,2'-anhydro ring opening strategy reported by Saneyoshi *et al.*,¹⁹ using excess trimethylsilylated alcohol as a nucleophile (masked anion). Mechanistically, this approach could be called as double activation, as the metal would activate the 2'-electrophilic carbon through co-ordination to the 4-carbonyl oxygen of **3/4** (electrophile activation) and the alkoxide/silylated as masked anion (nucleophile activation).

In the present design, we synthesize both the enantiomeric amino alcohol tethers from L-serine using a desymmetrization-like concept. The 2'-*O*-functionalization as reported by Saneyoshi *et al.*, for our designed molecule, we found that even TBS-ethers were equally reactive as TMS-ethers in the ring opening reaction of **3** (footnote, Table 1).²⁰ We reasoned that the 'O-Si' bond breaking of activated/protected silyl ethers (**8(S)**/**10(R)**) may not be involved in the rate determining step. Hence the 2'-electrophilic carbon activation in **3** using $\text{BF}_3 \cdot \text{OEt}_2$ must be the step that determines the reaction, rather than the alcohol partner. Therefore, conceptually, the addition of $\text{BF}_3 \cdot \text{OEt}_2$ to **3** would activate the 2'-electrophilic carbon. Subsequently, addition of the free alcohol may then provide the corresponding 2'-*O*-functionalized uridine.

We initially examined this concept during the synthesis of 2'-*O*-methoxyethyl- (MOE) and 2'-*O*-methoxypropyl- (MOP) controls, required for our studies. Interestingly, as anticipated, free alcohols 2-methoxyethanol and 3-methoxypropanol efficiently opened the

Chapter 2

2,2'-anhydrouridine ring in presence of $\text{BF}_3 \cdot \text{Et}_2\text{O}$, i.e., the 2'-*O*-functionalization was achieved with the activation of only the electrophile (Scheme 1).

Scheme 1 Efficient strategy for 2'-*O*-functionalization of pyrimidine-nucleosides

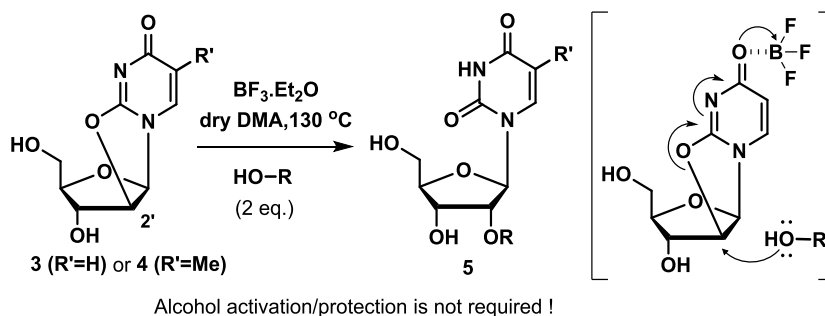


Table 1 An efficient 2'-*O*-functionalization of pyrimidine nucleosides with various alcohols*

S. No	HO-R	R'	Isolated yield (%)
1		H	70
2		H	69
3		H	64
4		H	62
5		Me	72
6		Me	60
7		Me	63

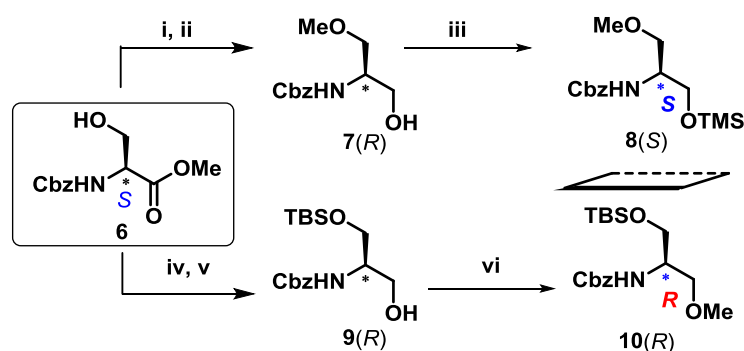
*Initially the TMS-ether and TBS-ether of alcohols (entry 3 and 4 respectively) were used in the ring opening reaction of **3**, when similar yields were obtained.

We extended this methodology to the synthesis of our branched 2'-*O*-*R/S*-AMP tethered uridine derivatives, using *R/S*-2-amino-3-methoxypropanol derivatives as

nucleophiles and obtained similar results (Table 1). We found that the free alcohols discussed above are also effective in ring opening of 2,2'-anhydrothymidine. Hence, we successfully simplified the strategy of 2'-*O*-functionalization of uridine and ribothymidine. The uridine and thymidine derivatives thus obtained could also be converted to the cytidine and 5-methylcytidine derivatives respectively by using the reported protocols.^{17, 19, 21} Consequently, we developed an efficient strategy for 2'-*O*-functionalization of pyrimidine nucleosides through our mono-activation strategy (of electrophile) rather than the reported double-activation strategies and using only two equivalents of the alcohol.^{18-19, 22}

2A.4 Synthesis of enantiomeric tethers from L-serine

Scheme 2 Synthesis of enantiomeric tethers from *N*-Cbz-L-serine methylester



Reagents and conditions: (i) Ag₂O, MeI, dry ACN, 12 h (ii) NaBH₄, MeOH 0 °C- rt, 8 h, 83% over two steps (iii) TMS-Cl, NEt₃, dry ACN, 1 h, 83% (iv) TBS-Cl, imidazole, dry DCM, 6 h (v) NaBH₄, MeOH, 0 °C- rt, 87% over two steps (vi) Ag₂O, MeI, dry ACN, 12 h, 91%.

Stereospecific syntheses of protected [*R/S*-(2-amino-3-methoxy)propanol] tethers (*R*-AMP and *S*-AMP) were achieved from L-serine by using a desymmetrization-like concept. The synthesis of protected [*S/R*-(2-amino-3-methoxy)propanol] **8(S)** and **10(R)** was achieved this starting from the same chiral source, *i.e.*, L-serine, by simply altering the sequence of reactions to get the individual enantiomeric tethers **8(S)** and **10(R)** (supported by chiral HPLC analysis, Experimental section) having suitable activation/protection, as shown in Scheme 2.

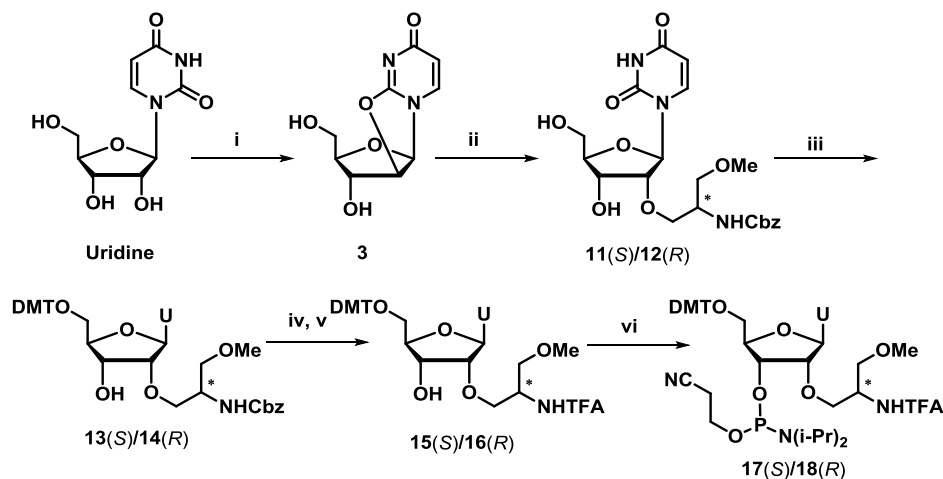
The *N*-Cbz-protected-L-serine methyl ester **6** was *O*-methylated using silver oxide and methyl iodide. After reduction with NaBH₄ in methanol, hydroxy compound **7(R)** was

obtained (83% overall yield) and was further activated as a silyl ether **8(S)** (83%) using trimethylsilylchloride (TMS-Cl). Alternatively, ester **6** was first protected as *tert*-butyldimethylsilyl ether (TBS) and then the ester was reduced to get alcohol **9(R)** (87%). Methylation of the free hydroxyl group in **9(R)** gave TBS protected ether **10(R)** in 91% yield.

2A.5 Synthesis of *R/S*-AMP phosphoramidite derivatives of uridine

Synthesis of appropriately protected *R*-AMP and *S*-AMP phosphoramidite derivatives of uridine were achieved starting from 2,2'-anhydrouridine **3**.

Scheme 3 Synthesis of *R/S*-AMP phosphoramidite derivatives of uridine



Reagents and conditions: (i) Ph_2CO_3 , NaHCO_3 , dry DMF, 130 °C (ii) **7(R)** or **8(S)** for **11(S)** and **10(R)** for **12(R)**, $\text{BF}_3 \cdot \text{Et}_2\text{O}$, dry DMA, 130 °C, 8-12 h, 60-68% (iii) DMT-Cl, cat.DMAP, dry pyridine, 6 h, 89% (iv) H_2 , Pd/C, 65 psi, MeOH, rt, 6 h (v) Ethyltrifluoroacetate, NEt_3 , MeOH, rt, 10 h, 86-88% over two steps (vi) 2-cyanoethyl-*N,N*-diisopropylchlorophosphine, DIPEA, dry DCM, 3 h, 88-92%.

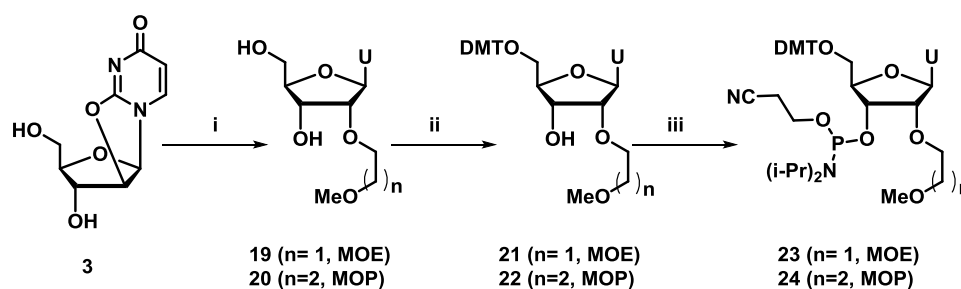
2, 2'-anhydrouridine **3** was prepared from the commercially available uridine, by heating with diphenyl carbonate in dry DMF at 130 °C using catalytic amount of NaHCO_3 . We used our strategy to get the desired 2, 2'-anhydro ring-opened products **11(S)** and **12(R)** in moderate yields (60% and 68% respectively) by employing 2 equivalents of **7(R)** and **10(R)**, Scheme 3. Compounds **11(S)** and **12(R)** were then protected as 5'-*O*-DMT derivatives to get **13(S)** and **14(R)** respectively (Scheme 3). The *N*-Cbz protecting group of

the 2'-alkoxy derivatives was removed under hydrogenation conditions using 10% Pd/C. The primary amino group was then protected to get the trifluoroacetyl derivatives **15(S)** and **16(R)**,^{15c} suitable for the solid-phase DNA synthesis protocols, Scheme 3. The phosphoramidite derivatives **17(S)** and **18(R)** were then prepared by phosphitylation using known method²³ starting from **15(S)** and **16(R)**, to be used in automated DNA synthesis.

2A.6 Synthesis of 2'-*O*-MOE and 2'-*O*-MOP controls

In order to compare the duplex stability of our *R/S*-AMP derived modified oligonucleotides, we synthesized phosphoramidite derivatives of 2'-*O*-methoxyethyl- (2'-*O*-MOE) (**23**) (a modification that is an important feature in the antisense drug Mipomersen, the only antisense drug currently in the market) and 2'-*O*-methoxypropyl- (2'-*O*-MOP) (**24**) uridine derivatives. The anhydrouridine **3** ring opening was effected with 2-methoxyethanol or 3-methoxypropanol to get the ring opened products **19** and **20** respectively. The 5'-OH groups of uridine derivatives **19** and **20** were protected as DMT-ethers, giving **21** and **22** which were further subjected to phosphitylation reaction using 2-cyanoethyl-*N,N*-diisopropylchlorophosphine to furnish the amidites **23** and **24** respectively.

Scheme 4 Synthesis of MOE and MOP phosphoramidite derivatives of uridine



Reagents and conditions: (i) 2-methoxyethanol/2-methoxypropanol, $\text{BF}_3 \cdot \text{Et}_2\text{O}$, dry DMA, 130 °C, 8-12 h, 69% (ii) DMT-Cl, cat.DMAP, dry pyridine, 6 h, 90-92% (iii) 2-cyanoethyl-*N,N*-diisopropylchlorophosphine, DIPEA, DCM, 3 h, 84-88%.

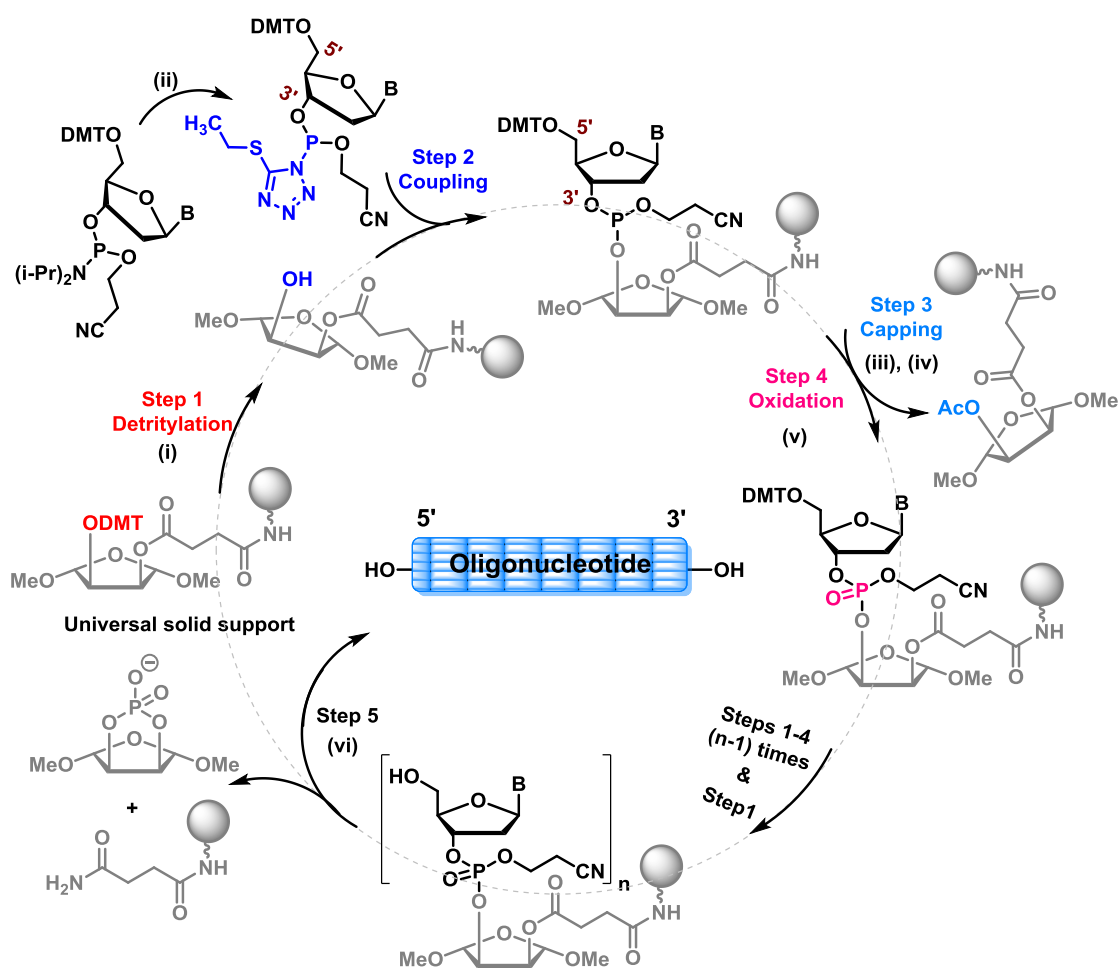
2A.7 Solid phase synthesis of oligonucleotide using phosphoramidite chemistry

Two types of solid supports are available on which oligonucleotide synthesis can be carried out. One is the standard support carrying a nucleosidic unit, which becomes the 3'-

Chapter 2

terminal residue, as the synthesis is done in 3'→5' direction. So, to use standard supports one requires four different supports for DNA synthesis carrying the individual canonical nucleobases and also four additional supports for RNA synthesis. The other type is the universal solid support, which carries an abasic sugar unit rather than the 3'-nucleoside unit. The advantage of using the universal support is that four different supports for DNA synthesis are not required and another major advantage is the possibility to synthesize 3'-end modified nucleic acids which is not possible using conventional standard supports.

Scheme 5 The sequence of chemical reactions involved in the solid phase synthesis of oligonucleotide on universal support



Reagents and conditions: (i) 3% TCA in DCM (ii) 0.25 M 5-(S-ethyltetrazole) in ACN (iii) Ac₂O/Py (iv) 10% N-methylimidazole in THF 0.1 M (v) I₂/Py/H₂O/THF (vi) aq. NH₃, 55 °C.

The A, G, C and T monomers were protected at their 5'-ends with a lipophilic, acid labile 4, 4'-dimethoxytrityl group and the exocyclic amines of nucleobases were protected with base labile protecting groups. The coupling of each monomer takes place *via* an elegant synthetic cycle (Scheme 5). Sequential repetition of the four steps in the cycle gave rise to the 3'-5' linked requisite oligonucleotide. In the post synthetic treatment, the synthesized oligonucleotide was cleaved from the support by aqueous ammonia treatment which also removed the inter-nucleoside β -cyanoethyl phosphate protection, the exocyclic amino protection of the bases and the 3'-end abasic sugar residue (as cyclic phosphate) to get the desired oligonucleotides with the free terminal 5' and 3'-OH groups. The crude oligomer was purified by RP-HPLC and characterized by MALDI-TOF mass spectrometry.

2A.8 Synthesis of modified oligonucleotides, characterization and UV- T_m studies

We chose a DNA sequence of biological relevance that was used in the splice correction assay developed by Kang *et al.*²⁴ The unmodified oligomers were synthesized using a Bioautomation MM4 DNA synthesizer and commercially available phosphoramidite building blocks. The phosphoramidites **17(S)** and **18(R)** were incorporated into sequences using increased coupling time (6 min) to yield the modified sequences. The site of the modified units in the sequences was decided so that these units were in the middle of the sequence (**ON2-ON5**, Table 2), were separated by 4-5 nucleosides (**ON6-ON9**, Table 2) or were at consecutive 3' end positions (**ON11-ON14**, Table 2).

The AONs were purified by HPLC subsequent to ammonia treatment and characterized by MALDI-TOF mass spectrometric analysis. Their purity was also re-checked by analytical HPLC and found to be >95% prior to their use in experiments. The thermal stability of DNA:RNA duplexes was evaluated using temperature dependent UV measurements (Figures 5 and 6). The melting studies revealed that the modified oligonucleotides **ON2-ON9** bind relatively stronger to cRNA than cDNA.

Chapter 2

Table 2 Modified DNA sequences, their MALDI-TOF mass analyses and biophysical evaluation by UV- T_m measurements^a

Sequences/Code ^b	Mass		T_m (°C)	T_m (°C)
	Calcd.	Obsd.	RNA ^c (ΔT_m)	DNA ^d (ΔT_m)
cctcttacctcagttaca ON1	5366.9	5367.2	56.6	54.7
cctcttacctcagt $\underline{U}^{\text{MOE}}$ aca ON2	5429.9	5430.0	55.8(-0.8)	52.7(-2.0)
cctcttacctcagt $\underline{U}^{\text{MOP}}$ aca ON3	5440.9	5474.1	55.5(-1.3)	52.7(-2.0)
cctcttacctcagt $\underline{U}^{\text{S-AMP}}$ aca ON4	5459.6	5460.2	56.2(-0.4)	54.7(0.0)
cctcttacctcagt $\underline{U}^{\text{R-AMP}}$ aca ON5	5459.6	5460.6	56.2(-0.4)	54.0(-0.7)
cctcttacc $\underline{U}^{\text{MOE}}$ cagtt $\underline{U}^{\text{MOE}}$ aca ON6	5490.9	5492.7	57.4(+0.8)	49.4(-5.3)
cctcttacc $\underline{U}^{\text{MOP}}$ cagtt $\underline{U}^{\text{MOP}}$ aca ON7	5515.0	5541.6	56.5(-0.1)	49.4(-5.3)
cctcttacc $\underline{U}^{\text{S-AMP}}$ cagtt $\underline{U}^{\text{S-AMP}}$ aca ON8	5549.0	5549.8	57.3(+0.7)	53.4(-1.3)
cctcttacc $\underline{U}^{\text{R-AMP}}$ cagtt $\underline{U}^{\text{R-AMP}}$ aca ON9	5549.0	5551.4	57.5(+0.9)	51.7(-3.0)
tttttttttttttt ON10	3588.1	3589.7		
tttttttttttt $\underline{U}^{\text{MOE}}$ $\underline{U}^{\text{MOE}}$ ON11	3708.9	3711.6		
tttttttttttt $\underline{U}^{\text{MOP}}$ $\underline{U}^{\text{MOP}}$ ON12	3734.6	3742.3		
tttttttttttt $\underline{U}^{\text{S-AMP}}$ $\underline{U}^{\text{S-AMP}}$ ON13	3767.0	3770.1		
tttttttttttt $\underline{U}^{\text{R-AMP}}$ $\underline{U}^{\text{R-AMP}}$ ON14	3767.0	3771.3		

^aUV- T_m values were measured by using 1 μ M sequences with 1 μ M cDNA/cRNA in sodium phosphate buffer (0.01M, pH 7.2) containing 100 mM NaCl and are averages of three independent experiments. (Accuracy is ± 0.5 °C). ^bThe lower case letters indicate unmodified DNA, $\underline{U}^{\text{MOE}}$ / $\underline{U}^{\text{MOP}}$ denotes 2'-*O*-methoxyethyl-/2'-*O*-methoxypropyl uridine, $\underline{U}^{\text{S-AMP}}$ / $\underline{U}^{\text{R-AMP}}$ denotes 2'-*O*-[*S*-(2-amino-3-methoxy)propyl]/2'-*O*-[*R*-(2-amino-3-methoxy)propyl] uridine derivatives respectively. ^c5'uguaacugagguaagagg was the complementary RNA sequence. ^d5'tgtaactgaggaagagg was the complementary DNA sequence.

The chirality in the AMP tether of the modified oligonucleotides did not have much effect when binding to cRNA, but it did have a slight effect when binding to cDNA (**ON4** vs **ON5** and **ON8** vs **ON9**). When two units of $\underline{U}^{\text{S/R-AMP}}$ were incorporated in oligonucleotides (**ON8/ON9**), there was a slight increase in the duplex stability with the

cRNA, whereas the duplex stability decreased with cDNA (Table 2). It is also noteworthy that the duplex stability of the oligonucleotides **ON4-ON5** and **ON8-ON9** containing single and double modifications ($U^{S/R-AMP}$) with cRNA/cDNA were comparable to U^{MOE} containing oligonucleotides **ON2** and **ON6**, respectively (Table 2). We examined the *S*-type to *N*-type conformational equilibria for 2'-*O*-MOE-uridine (**19**) with **11(S)** and **12(R)** and found that they are comparable (%*S*= 56%, 46% and 55% for **19**, **11(S)** and **12(R)** respectively, Table 3, Experimental section).

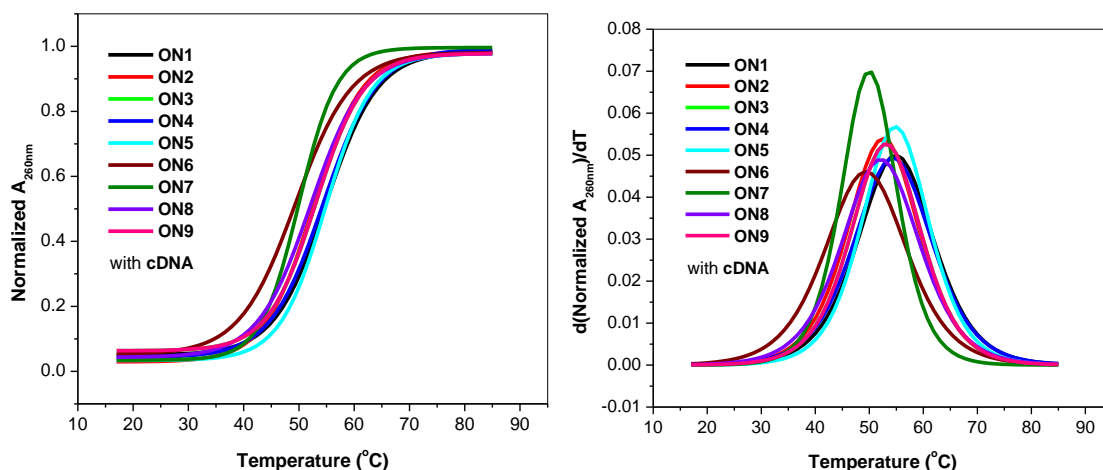


Figure 5 UV melting profiles (a) and the first order derivatives (b) of complexes of **ON1-ON9** with cDNA.

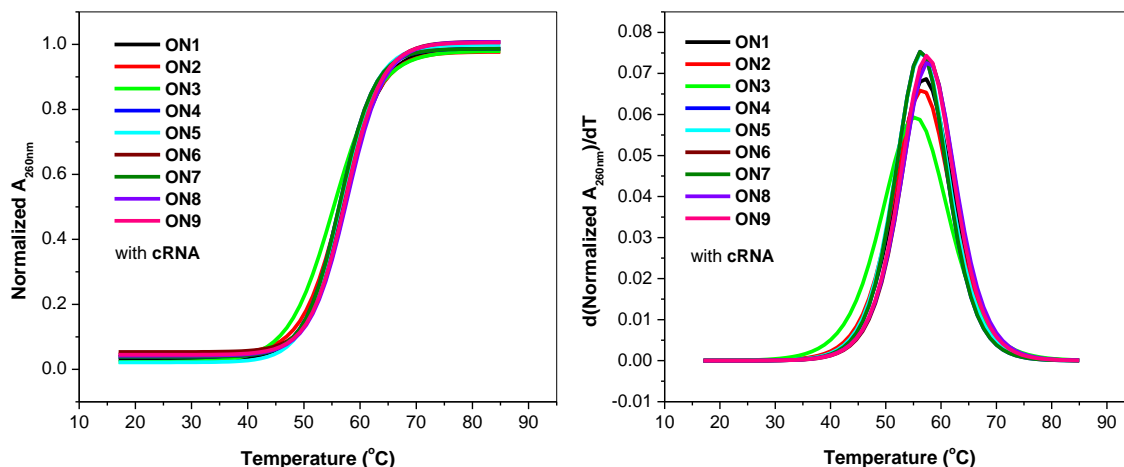


Figure 6 UV melting profiles (a) and the first order derivatives (b) of complexes of **ON1-ON9** with cRNA.

This observation could be attributed to the vicinal electronegative substituents at C2', namely, O4'- and O2'- and the 2-amino- and 3-methoxy- groups on the propyl chain at the 2'-*O*- position of uridine which would be expected to maintain the *gauche* orientations with respect to the O4'-oxygen to induce conformational change and hydration in the minor groove as is known for 2'-*O*-MOE derivatives.^{4a, 7}

2A.9 SVPD digestion of the modified oligonucleotides

To investigate the synergistic role of *S/R*-AMP tethers, we further explored the *in vitro* studies by examining the protection offered by the U^{*S*-AMP} and U^{*R*-AMP} tethers against hydrolytic cleavage in the newly synthesized oligomers and compared the results with the U^{MOE} tether.¹⁻² 2'-*O*-MOE is being developed as an antisense therapeutic and is also finding newer applications in the siRNA approach where only two terminal units are modified with 2'-*O*-MOE.³ We therefore followed literature precedence and synthesized a homothyminyl sequence^{15e, 18b, 22a, 25} **ON10** and modified it at the 3'-terminus with two consecutive units of U^{MOE}, U^{MOP}, U^{*S*-AMP} or U^{*R*-AMP} to get sequences **ON11**, **ON12**, **ON13** and **ON14** respectively (Figure 7). We digested these sequences with SVPD enzyme under conditions reported earlier.^{15e, 25a} The products of the digestion were analyzed by RP-HPLC and the percent intact AON was plotted against time to understand the degradation pattern for all the oligomers (Figure 7). The results clearly indicate the dependence of enzyme digestion on steric bulk and/or the presence of protonable amino group as charged species, as both the oligomers containing U^{*S*-AMP} **ON13** and U^{*R*-AMP} **ON14** were found to be much stabler than the unmodified **ON10**, U^{MOE}-modified **ON11** or U^{MOP}-modified **ON12** sequences. The stereochemical influence of the 2'-*O*-AMP tethers in U^{*S*-AMP} (**ON13**) and U^{*R*-AMP} (**ON14**) AONs was clearly reflected in the SVPD digestion experiment (Figure 7). The 2'-*O*-MOE- and 2'-*O*-MOP-AONs **ON11** and **ON12** were almost completely digested by SVPD at the end of 1 h. At the same time, almost 90% of 2'-*O*-*R*-AMP-AON (**ON14**) and 70% 2'-*O*-*S*-AMP-AON (**ON13**) remained intact and were effectively able to resist the hydrolytic cleavage. Furthermore, ~42 % of 2'-*O*-*R*-AMP-tethered **ON14** was still intact after 4 hours and was almost thrice as stable as the *S*-stereoisomer.

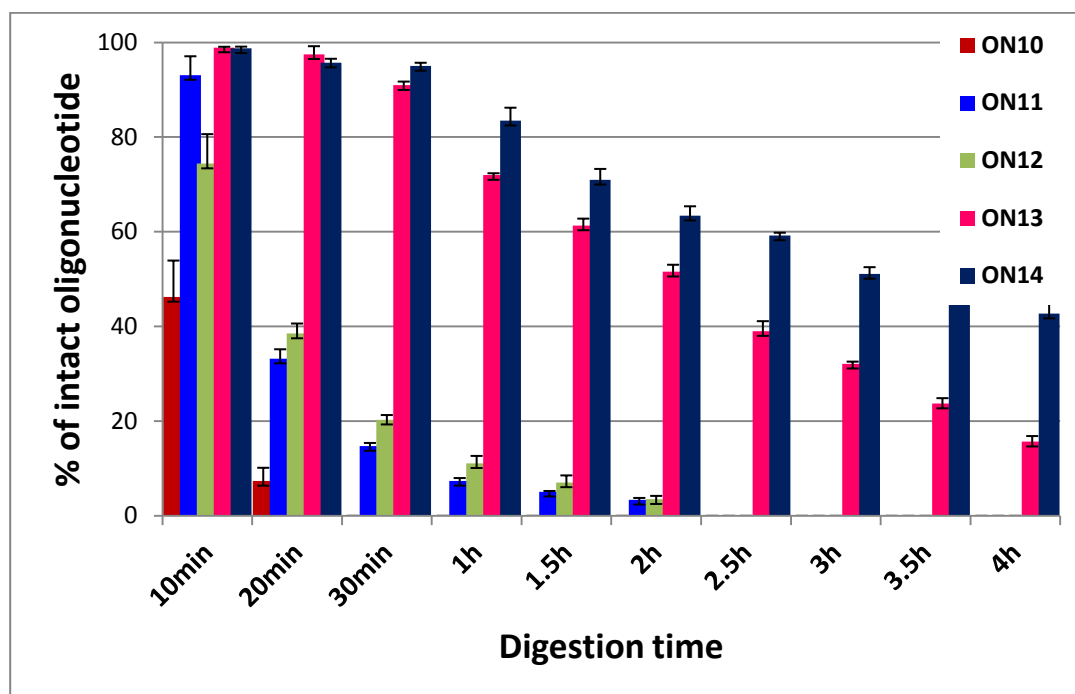


Figure 7 Preferential protection of 3'-end modified AONs, with 2'-*O*-*S*-AMP- (**ON13**) and 2'-*O*-*R*-AMP- substituents (**ON14**) compared to native DNA (**ON10**), 2'-*O*-MOE-modified DNA (**ON11**) and 2'-*O*-MOP-modified DNA (**ON12**). The experiment was done in duplicate and the error bars (standard deviations) are indicated. (Please see the experimental section for the conditions used).

Introduction of chirality in the minor groove as in the case of 2'-*O*-*R*-AMP and 2'-*O*-*S*-AMP, did indeed allow differential protection by one of the isomers while interacting with the active site of the enzyme,¹⁰ implying that this interaction could be energetically different for the two isomers, compared to the achiral, acyclic, alkyl substitutions such as 2'-*O*-MOE or 2'-*O*-AE/AP.^{4a, 7} It could be presumed that the projection of the amino substituent in the minor groove while interacting with the chiral environment of the enzyme⁷ would play a crucial role towards this result. The chirality of our new compounds thus allowed preferential protection against nucleases synergistically while retaining the duplex stability, which in turn would improve the therapeutic potency of 2'-*O*-*R*-AMP-AONs. A similar chirality effect on protection of AONs against digestive enzymes such as SVPD was also observed when fully modified, stereorandom phosphorothioate oligonucleotides were synthesized in chirally pure form,²⁶ but in that case the duplex stability was found to be compromised.

2A.10 Conclusions

- ✓ The 2'-*O*-MOE modification has been transformed into a synergy of chirality, positive charge and hydration module, all in one, in the 2'-*O*-*R/S*-AMP-modified AONs.
- ✓ An efficient synthetic route for 2'-*O*-functionalization of pyrimidine nucleosides through a new mono-activation strategy was accomplished.
- ✓ The synthesis of 2'-*O*-*R/S*-AMP enantiomeric tethers was achieved from the same chiral source L-serine by using a desymmetrization-like concept.
- ✓ The synergistic contribution of 2'-*O*-MOE and chiral amino substituents preserved the DNA:RNA duplex stability and their stereochemistry-dependent enzyme digestion was found to be much more effective compared to 2'-*O*-MOE alone.
- ✓ The demonstrated preserved DNA:RNA duplex stability (like 2'-*O*-MOE, a modification that is being evaluated in clinical trials for several diseases) and the improved stability against hydrolytic enzymes are desired and necessary properties to improve the therapeutic value of any drug candidate in development.
- ✓ Thus, our designed 2'-*O*-*R/S*-AMP tethered oligonucleotides have high potential for applications in antisense and siRNA therapeutic development.

2A.11 Experimental Section

General information

All the non-aqueous reactions were carried out under the inert atmosphere of Nitrogen/Argon and the chemicals used were of laboratory or analytical grade. All solvents used were dried and distilled according to standard protocols. TLCs were carried out on pre-coated silica gel GF254 sheets (Merck 5554). Column chromatographic separations were performed using silica gel 100-200 mesh (Merck) or 230- 400 mesh (Merck) and using the solvent systems EtOAc/Petroleum ether or MeOH/DCM. ^1H , ^{13}C and ^{31}P NMR spectra were obtained using Bruker AC-200, AC-400 or AC-500 NMR spectrometers. The chemical shifts are reported in delta (δ) values and referred to internal standard TMS for ^1H . HRMS mass spectra were recorded on a Thermo Scientific Q-Exactive LC-MS/MS mass spectrometer. DNA and modified oligonucleotides were synthesized on a Bioautomation Mer-Made 4 DNA synthesizer using standard β -cyanoethyl phosphoramidite chemistry. Complementary RNA sequence for thermal denaturation experiments (5'-UGUAACUGAGGUAAGAGG-3') was purchased from Sigma-Aldrich. Oligomers were analyzed and purified by RP HPLC (Waters Delta 600e quaternary solvent delivery system and 2998 photodiode array detector and Empower2 chromatography software) on a C18 column using an increasing gradient of acetonitrile in 0.1N triethylammonium acetate, pH 7.0, and characterized by MALDI-TOF mass spectrometry. The MALDI-TOF spectra were recorded on Voyager-De-STR (Applied Biosystems) MALDI-TOF or AB Sciex TOF/TOFTM series explorerTM 72085 instrument and the matrix used for analysis was THAP (2, 4, 6-trihydroxyacetophenone). Thermal denaturation experiments were performed on Varian Cary-300 UV-Vis spectrophotometer fitted with a peltier-controlled temperature programmer and a water circulator, at the temperature ramping rate 0.5 °C/min and the absorbance was recorded for every 0.5 °C rise in temperature.

Chapter 2

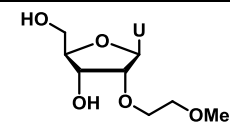
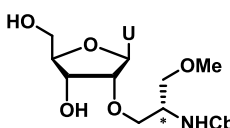
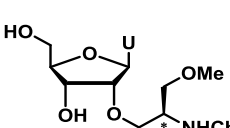
SVPD digestion: Stability of the oligonucleotides **ON10-ON14** towards exo-nucleases SVPD (snake venom phosphodiesterase) was analyzed by RP-HPLC. 7.5 μ M of oligonucleotide in 300 μ L of Tris-HCl buffer (pH=7.5, 10mM Tris-HCl, 8mM MgCl₂) was incubated at 37 °C for 15 minutes. SVPD 5 ng/ μ L was added to the oligonucleotide at 37 °C and aliquots of 20 μ L were removed at time intervals of 0, 2, 5, 10, 20, 30, 60, 90, 120, 150, 180, 210 and 240 minutes. Aliquots were kept at 90 °C for 2 minutes and analysed by RP-HPLC using an increasing gradient of acetonitrile in triethylammonium acetate (A: 5% acetonitrile and B: 30% acetonitrile in 0.1N triethylammonium acetate, pH 7.0). The percent intact oligonucleotide (based on the peak area) was plotted against the time.

Conformational analysis of **19**, **11(S)** and **12(R)**:

Based on ¹H NMR coupling constants of H1' and H2', the %S conformation was calculated using the simple equation (1),²⁷ as shown in Table 3.

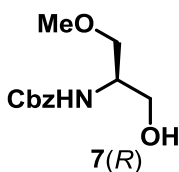
$$\%S = 100 \times (J_{H1'-H2'} - 1)/6.9 \quad (1)$$

Table 3 %S calculations at the monomer level

Chemical Structure	$J_{H1'-H2'}(\text{Hz})$	$\%S=(J_{H1'-H2'}-1) \times 100/ 6.9$
 19	4.88	56.2
 11(S)	4.15	45.7
 12(R)	4.77	54.6

(R)-benzyl (1-hydroxy-3-methoxypropan-2-yl)carbamate, 7(R):

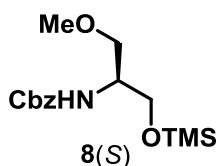
N-Cbz protected-L-serine-methyl ester **6** (39.5 mmol, 10 g) was dissolved in dry acetonitrile (250 mL), followed by the addition of Ag₂O (98.8 mmol, 22.8 g) and MeI (197.5 mmol, 12.7 mL) and the reaction mixture was vigorously stirred at room temperature for 12 hours. The reaction mixture was filtered and the filtrate was concentrated under reduced pressure to give the *O*-methylated derivative of **6**. The crude colourless residue was dissolved in MeOH (500 mL) and NaBH₄ (150 mmol, 5.6 g) was added fraction-wise at 0 °C for a period of 1 hour and the mixture was kept stirring at room temperature for another 6-8 hours. Excess NaBH₄ was quenched with saturated NH₄Cl solution, followed by the removal of MeOH under reduced pressure. The crude reaction mixture was extracted with EtOAc. The organic extract was washed with brine solution and dried over anhydrous sodium sulphate. EtOAc was removed under reduced pressure to give the crude product, that was purified by column chromatography (eluted in 20% EtOAc in petroleum ether) to yield **7(R)** as a colourless liquid in 83% yield (7.8 g) over two steps. ¹H NMR (CDCl₃, 200 MHz): δ 2.53 (bs, 1H), 3.35 (s, 3H), 3.55-3.86 (m, 5H), 5.11 (s, 2H), 5.43 (bs, 1H), 7.36-7.37 (m, 5H); ¹³C NMR (CDCl₃, 100 MHz): δ 51.8, 58.7, 62.1, 66.4, 71.7, 127.7, 128.1, 136.0, 156.2; HRMS (EI): Mass calculated for C₁₂H₁₈O₄N (M+H) 240.1230, found 240.1229.



7(R)

(S)-benzyl (1-methoxy-3-((trimethylsilyl)oxy)propan-2-yl)carbamate, 8(S):

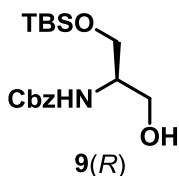
To a stirred solution of **7(R)** (16.7 mmol, 4 g) and NEt₃ (83.5 mmol, 11.7 mL) in dry acetonitrile (70 mL), TMS-Cl (25.1 mmol, 3.1 mL) was added and stirring was continued for another 1 hour. The reaction mixture was diluted with EtOAc. Water wash and brine wash were given to the organic layer. The organic layer was dried over anhydrous sodium sulphate, solvents removed in *vacuo* and the crude compound was purified by column chromatography (eluted in 5% EtOAc in petroleum ether) to give **8(S)** as a colourless liquid in 83% yield (4.3 g). ¹H NMR (CDCl₃, 200 MHz): δ 0.10 (s, 9H), 3.32 (s, 3H), 3.35-



3.39(m, 1H), 3.47-3.50(m, 1H), 3.56-3.60(m, 1H), 3.68-3.73(m, 1H), 3.84(m, 1H), 5.09 (s, 2H), 5.27 (bs, 1H), 7.29-7.34 (m, 5H); ^{13}C NMR (CDCl₃, 100 MHz): δ 0.8, 51.3, 58.6, 60.6, 66.4, 70.3, 127.8, 128.2, 136.3, 155.7; **HRMS (EI)**: Mass calculated for C₁₅H₂₆O₄NSi (M+H) 312.1626, found 312.1622.

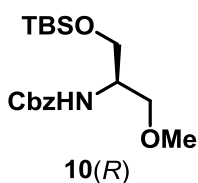
(R)-benzyl (1-((tert-butyldimethylsilyl)oxy)-3-hydroxypropan-2-yl)carbamate, 9(R):

N-Cbz protected-L-serine-methyl ester **6** (39.5 mmol, 10 g) was dissolved in dry DCM (200 mL), followed by the addition of imidazole (98.8 mmol, 6.7 g) and TBS-Cl (47.4 mmol, 7.1 g). The reaction mixture was diluted with DCM and the DCM layer was washed with water and brine solution. The organic layer was dried over anhydrous sodium sulphate and solvent removed in *vacuo* to



give the crude TBS protected ester, which was directly subjected to NaBH₄ reduction. The colourless residue was dissolved in methanol (500 mL) and NaBH₄ (150 mmol, 5.6 g) was added fraction-wise at 0 °C for a period of 1 hour and then stirring continued at room temperature for another 6 hours. Excess NaBH₄ was quenched with saturated NH₄Cl solution, followed by the removal of MeOH under reduced pressure and the crude compound was extracted with EtOAc. The organic extract was washed with brine and dried over anhydrous sodium sulphate. EtOAc was removed under reduced pressure to give the crude product and was purified by column chromatography (eluted in 15% EtOAc in petroleum ether) to yield **9(R)** as a colourless liquid in 87% yield (11.6 g) over two steps. ^1H NMR (CDCl₃, 200 MHz): δ 0.05 (s, 6H), 0.88 (s, 9H), 2.64 (bs, 1H), 3.66-3.84 (m, 5H), 5.11 (s, 2H), 5.38-5.41 (m, 1H), 7.35-7.38 (m, 5H); ^{13}C NMR (CDCl₃, 50 MHz): δ -5.6, 18.1, 25.7, 53.0, 63.0, 63.4, 66.7, 128.0, 128.4, 136.2, 158.2; **HRMS (EI)**: Mass calculated for C₁₇H₃₀O₄NSi (M+H) 340.1939, found 340.1945.

(R)-benzyl (1-((tert-butyldimethylsilyl)oxy)-3-methoxypropan-2-yl)carbamate, 10(R):



To a stirred solution of **9(R)** (29.4 mmol, 10 g) and MeI (147.4 mmol, 9.5 mL), Ag₂O (73.5 mmol, 16.9 g) was added and the reaction mixture was vigorously stirred at room temperature for 12 hours. The reaction mixture

was filtered and the filtrate was concentrated under reduced pressure. The crude compound was purified by column chromatography (eluted in 5% EtOAc in petroleum ether) to give **10(R)** as a colourless liquid in 91% (9.4 g) yield. **¹H NMR** (CDCl₃, 200 MHz): δ 0.05 (s, 6H), 0.88 (s, 9H), 3.33 (s, 3H), 3.37-3.77(m, 4H), 3.78-3.86(m, 1H), 5.10 (s, 2H), 7.35-7.38 (m, 5H); **¹³C NMR** (CDCl₃, 100 MHz): δ -5.6, 18.1, 25.7, 51.5, 58.8, 61.3, 66.6, 70.4, 128.0, 128.4, 136.4, 155.9; **HRMS (EI)**: Mass calculated for C₁₈H₃₂O₄NSi (M+H) 354.2095, found 354.2089.

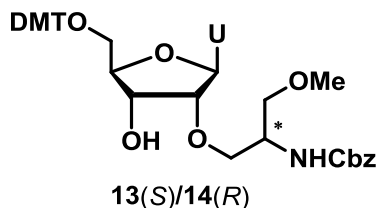
A typical procedure for 2'-O-functionalization of 2,2'-anhydrouridine:

Desiccated 2,2'-anhydrouridine **3** (6.6 mmol, 1.5 g) was dissolved in dry DMA (10 mL) followed by the addition of BF₃.OEt₂ (10 mmol, 1.2 mL) under argon atmosphere. After 2 minutes, the activated/protected silyl ethers **8(S)/ 10(R)** (13.2 mmol) were added and stirred at 130 °C for 8-12 h. DMA was removed partially in *vacuo*, followed by the dilution with MeOH and the reaction mixture was column purified (eluted in 5% MeOH in DCM) to furnish **11(S)/12(R)** as a white solid in 60-68% yield. **11(S)** **¹H NMR** (DMSO-d₆, 200 MHz): δ 3.21 (s, 3H), 3.49-3.63(m, 5H), 3.82-3.9 (m, 3H), 4.03-4.10 (m, 1H), 5.01-5.05 (m, 3H), 5.14 (t, 1H, J= 4.69, 9.75Hz), 5.60-5.65 (m, 1H), 5.79 (d, 1H, J= 4.15Hz), 7.35 (m, 5H), 7.92 (d, 1H, J= 7.76Hz); **¹³C NMR** (DMSO-d₆, 200 MHz): δ 50.0, 58.1, 59.9, 65.3, 68.1, 69.1, 70.9, 81.5, 84.4, 86.1, 101.6, 127.7, 128.3, 136.9, 140.3, 150.3, 155.8, 163.1; **HRMS (EI)**: Mass calculated for C₂₁H₂₇O₉N₃Na (M+Na) 488.1640, found 488.1630. **12(R)** **¹H NMR** (DMSO-d₆, 400 MHz): δ 3.22 (s, 3H), 3.51-3.68(m, 5H), 3.78-3.93 (m, 3H), 4.09-4.13 (m, 1H), 4.97-5.051 (m, 3H), 5.16 (t, 1H, J= 4.77, 9.54Hz), 5.64-5.66 (m, 1H), 5.84 (d, 1H, J= 4.77Hz), 7.36 (m, 5H), 7.92 (d, 1H, J= 8.03Hz); **¹³C NMR** (DMSO-d₆, 100 MHz): δ 50.3, 58.4, 60.6, 65.5, 68.6, 69.3, 71.4, 81.8, 85.0, 86.3, 102.0, 128.0, 128.5, 137.2, 140.6, 150.7, 156.0, 163.4; **HRMS (EI)**: Mass calculated for C₂₁H₂₈O₉N₃ (M+H) 466.1820, found 466.1822.

5'-O-DMT-2'-O-S/R-(N-benzoyloxycarbonyl)AMP uridine, 13(S)/14(R):

11(S)/12(R) (3.44 mmol, 1.6 g) was dissolved in dry pyridine (10 mL) and DMT-Cl (3.61 mmol, 1.22 g) and catalytic amount of DMAP (~20 mg) were added. The reaction mixture

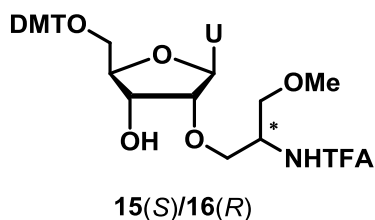
was kept for stirring at room temperature for 5-6 hours. Pyridine was removed under reduced pressure and the residue was diluted with EtOAc. 10% aqueous NaHCO₃, water



and brine solution washes were given to the organic layer. The organic layer was dried over anhydrous Na₂SO₄ and concentrated to dryness. The crude compound was column purified (eluted in 70% EtOAc in petroleum ether) to provide **13(S)/14(R)** as a white foam in 89% yield. **13(S)** ¹H

NMR (CDCl₃, 200 MHz): δ 3.33 (s, 3H), 3.42-3.54 (m, 5H), 3.8 (s, 6H), 3.91-4.04 (m, 4H), 4.44 (bs, 1H), 5.09 (s, 2H), 5.23 (d, 1H, J= 8.21Hz), 5.32 (d, 1H, J= 7.96Hz), 5.89 (s, 1H), 6.83-6.87 (m, 4H), 7.31-7.40 (m, 14H), 8.02 (d, 1H, J= 8.08Hz), 8.26 (bs, 1H); ¹³C **NMR** (CDCl₃, 125 MHz): δ 50.1, 55.0, 58.9, 61.0, 66.6, 68.2, 70.9, 71.4, 82.8, 82.9, 86.8, 87.5, 101.8, 113.1, 126.9, 127.8, 128.0, 128.3, 129.9, 130.0, 134.9, 135.2, 136.2, 139.9, 144.2, 150.2, 156.6, 158.4, 158.5, 163.8; **HRMS (EI)**: Mass calculated for C₄₂H₄₅O₁₁N₃Na (M+Na) 790.2946, found 790.2931. **14(R)** ¹H **NMR** (CDCl₃, 400 MHz): δ 3.31 (s, 3H), 3.50-3.52 (m, 5H), 3.76 (m, 6H), 3.90-3.91 (m, 1H), 4.02-4.04 (m, 3H), 4.39-4.42 (m, 1H), 5.07-5.08 (m, 2H), 5.27 (d, 1H, J= 8.24Hz), 5.64 (d, 1H, J= 8.24Hz), 5.92 (d, 1H, J= 0.92Hz), 6.82-6.84 (m, 4H), 7.27-7.39 (m, 14H), 7.96 (d, 1H, J= 8.24Hz); ¹³C **NMR** (DMSO-d₆, 100 MHz): δ 50.2, 55.0, 58.9, 61.2, 66.7, 68.3, 70.9, 71.5, 83.1, 83.2, 86.8, 87.2, 101.9, 113.1, 127.0, 127.8, 128.0, 128.3, 129.9, 130.0, 134.9, 135.1, 136.1, 139.8, 144.2, 150.3, 156.2, 158.50, 158.54, 163.7; **HRMS (EI)**: Mass calculated for C₄₂H₄₅O₁₁N₃Na (M+Na) 790.2946, found 790.2946.

5'-O-DMT-2'-O-S/R-(N-trifluoroacetyl)AMP uridine, **15(S)/16(R)**:



The 5'-DMT protected 2'-O-functionalized uridine derivative **13(S)/14(R)** (2.3 mmol, 1.8 g) was dissolved in MeOH (10 mL) followed by the addition of 10% Pd-C (10% w/w, 0.18 g). Then reaction mixture was subjected to catalytic hydrogenation at 65 psi of hydrogen pressure for 6 h. After TLC analysis, the reaction mixture was filtered over celite and the removal of

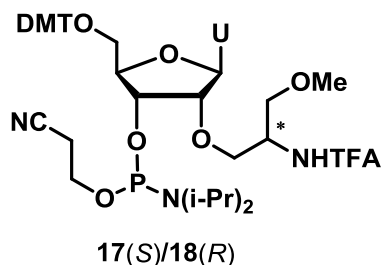
methanol in *vacuo* gave the free amine. Without further purification, the amine was subjected to trifluoroacetyl protection. To the crude amine (2.2 mmol, 1.4 g) dissolved in MeOH (15 mL), NEt₃ (3.3mmol, 0.46 mL) was added. Ethyltrifluoroacetate (2.76 mmol, 0.33 mL) was added to the reaction mixture and stirred at room temperature for 8-10 h. MeOH was removed on rota evaporator and the reaction mixture was diluted with ethylacetate. The organic layer was washed with water and 5% aq. NaHCO₃ and the organic layer was dried over anhydrous Na₂SO₄, concentrated *in vacuo*. The crude compound was column-purified to furnish **15(S)/16(R)** as a white foam in 86-88% yield (eluted in 55% EtOAc in petroleum ether). **15(S)** ¹H NMR (CDCl₃, 500 MHz): δ 3.40 (s, 3H), 3.52-3.63 (m, 5H), 3.81 (s, 6H), 3.82-3.85 (m,1H), 3.93 (dd,1H, *J*= 1.8 and 4.9Hz), 4.03-4.06 (m, 1H), 4.09 (dd,1H, *J*= 4.8 and 10.0Hz), 4.33-4.37 (m, 1H), 4.42-4.46 (m,1H), 5.30 (dd, 1H, *J*_{5,6}= 2.1 and 8.24Hz), 5.92 (d, 1H, *J*= 1.5Hz), 6.85-6.86 (m, 4H), 7.25-7.39 (m, 9H), 8.00 (d, 1H, *J*_{5,6}= 8.24Hz); ¹³C NMR (CDCl₃, 500 MHz): δ 49.43, 55.2,59.3, 61.2, 68.5, 70.1, 70.6, 83.4, 83.6, 87.1, 87.5, 102.3, 113.2, 113.3, 127.2, 128.0, 128.1, 130.0, 130.1, 135.0, 135.2, 139.7, 144.2, 150.2, 158.7, 162.7; **HRMS (EI)**: Mass calculated for C₃₆H₃₈O₁₀N₃F₃Na (M+Na) 752.2402, found 752.2391. **16(R)** ¹H NMR (DMSO-d₆, 400 MHz): δ 3.24 (s, 3H), 3.43-3.45 (m, 3H), 3.64-3.68 (m, 2H), 3.74 (s, 6H), 3.83 (m,1H), 3.94-3.99 (m, 3H), 4.16-4.19 (m, 3H), 5.17 (d, 1H, *J*= 7.50Hz), 5.27 (dd, 1H, *J*_{5,6}= 2.0 and 8.17Hz), 5.77 (d, 1H, *J*= 2.75Hz), 6.89-6.91 (m, 4H), 7.24-7.39 (m, 9H), 7.70 (d, 1H, *J*_{5,6}= 8.19Hz); ¹³C NMR (CDCl₃, 500 MHz): δ 49.44, 55.2, 59.2, 60.9, 68.4, 69.3, 70.4, 83.1, 83.4, 87.1, 87.8, 102.2, 113.1, 113.3, 127.0, 128.0, 128.1, 130.0, 131.0, 135.0, 135.2, 139.7, 144.3, 150.3, 158.6, 158.7; **HRMS (EI)**: Mass calculated for C₃₆H₃₈O₁₀N₃F₃Na (M+Na) 752.2402, found 752.2405.

General procedure followed for the synthesis of phosphoramidite derivatives, 17(S)/18(R):

To the compound **15(S)/16(R)** (0.68 mmol, 500mg) dissolved in dry DCM (10 mL), DIPEA (1.7 mmol, 0.29mL) was added. 2-cyanoethyl-*N,N*-diisopropyl-chloro phosphine (0.81 mmol, 0.18 mL) was added to the reaction mixture at 0 °C and stirring continued at

Chapter 2

room temperature for 3 hours. The contents were diluted with DCM and washed with 5%



NaHCO₃ solution. The organic phase was dried over anhydrous sodium sulphate and concentrated to white foam.

The residue was re-dissolved in DCM and the compound was precipitated with n-hexane to obtain corresponding phosphoramidite derivatives in 88-92 % yield.

Phosphoramidite **17(S)**: ³¹P NMR (Acetonitrile, D₂O as external standard, 400 MHz): δ 149.22, 149.68; **HRMS (EI)**: Mass calculated for C₄₅H₅₅O₁₁N₅F₃PNa (M+Na) 952.3480, found 952.3516.

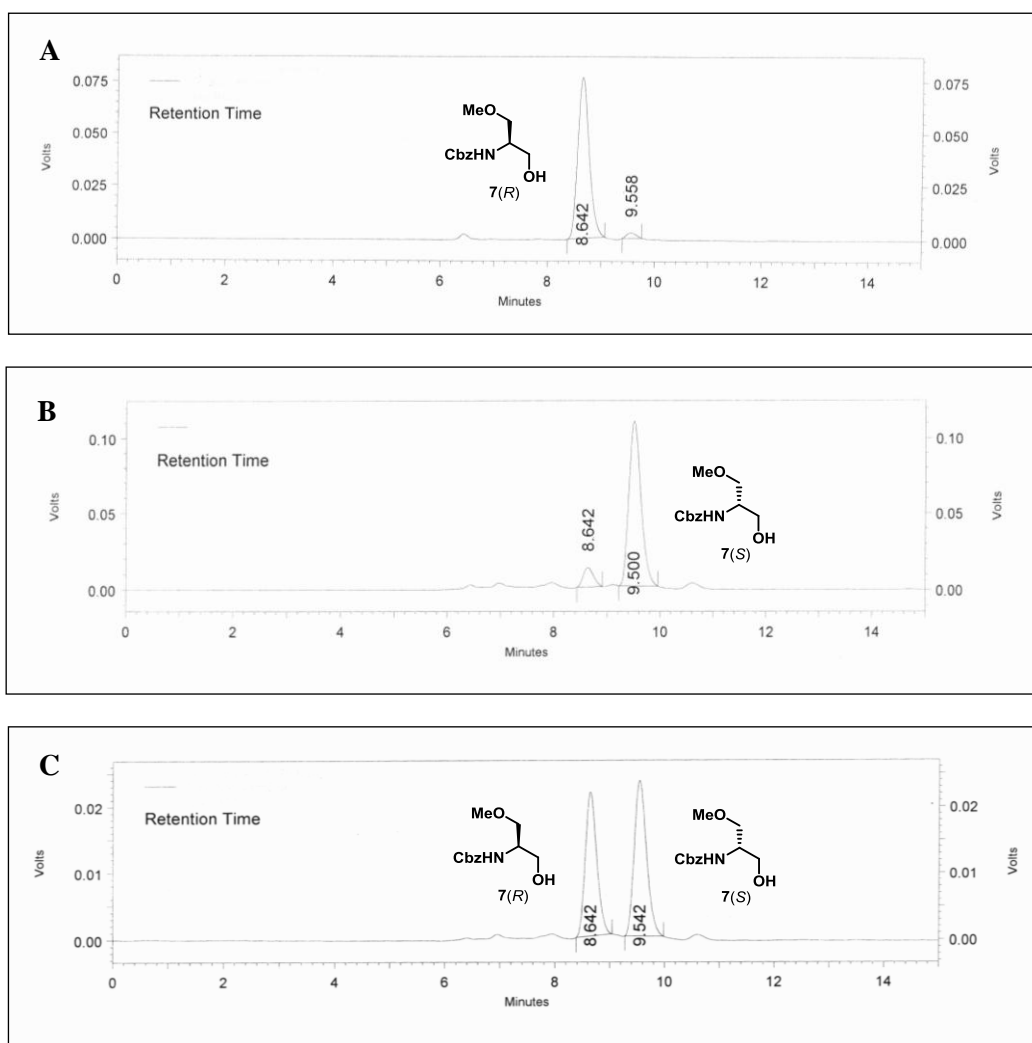
Phosphoramidite **18(R)**: ³¹P NMR (Acetonitrile, D₂O as external standard, 400 MHz): δ 149.52, 150.62; **HRMS (EI)**: Mass calculated for C₄₅H₅₆O₁₁N₅F₃P (M+H) 930.3661, found 930.3674.

2A.12 Appendix

Compounds - Spectral data	Page No.
Chiral HPLC chromatograms of enantiomeric tethers from L-serine	70
7- ^1H , ^{13}C NMR & DEPT	71
8- ^1H , ^{13}C NMR & DEPT	72
9- ^1H , ^{13}C NMR & DEPT	73
10- ^1H , ^{13}C NMR & DEPT	74
11- ^1H , ^{13}C NMR & DEPT	75
12- ^1H , ^{13}C NMR & DEPT	76
13- ^1H , ^{13}C NMR & DEPT	77
14- ^1H , ^{13}C NMR & DEPT	78
15- ^1H , ^{13}C NMR & DEPT	79
16- ^1H , ^{13}C NMR & DEPT	80
17- ^{31}P , 18- ^{31}P	81
7, 8- HRMS	82
9, 10- HRMS	83
11, 12- HRMS	84
13, 14- HRMS	85
15, 16- HRMS	86
17, 18- HRMS	87
ON1, ON2- MALDI-TOF spectra	88
ON3, ON4- MALDI-TOF spectra	89
ON5, ON7- MALDI-TOF spectra	90
ON8, ON9- MALDI-TOF spectra	91
ON10, ON11- MALDI-TOF spectra	92
ON12, ON13- MALDI-TOF spectra	93
ON14- MALDI-TOF spectra	94
ON10-ON14, SVPD digestion- HPLC chromatograms overlay	95-96

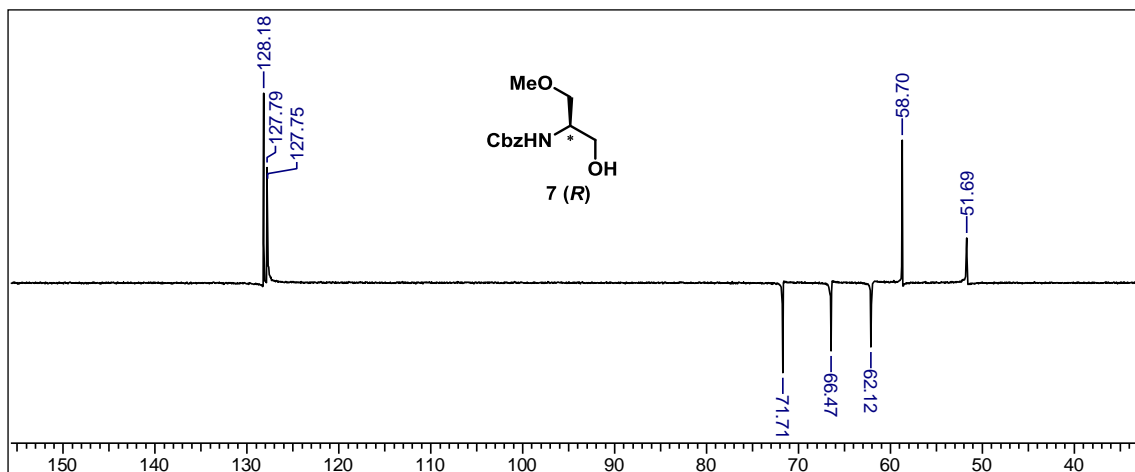
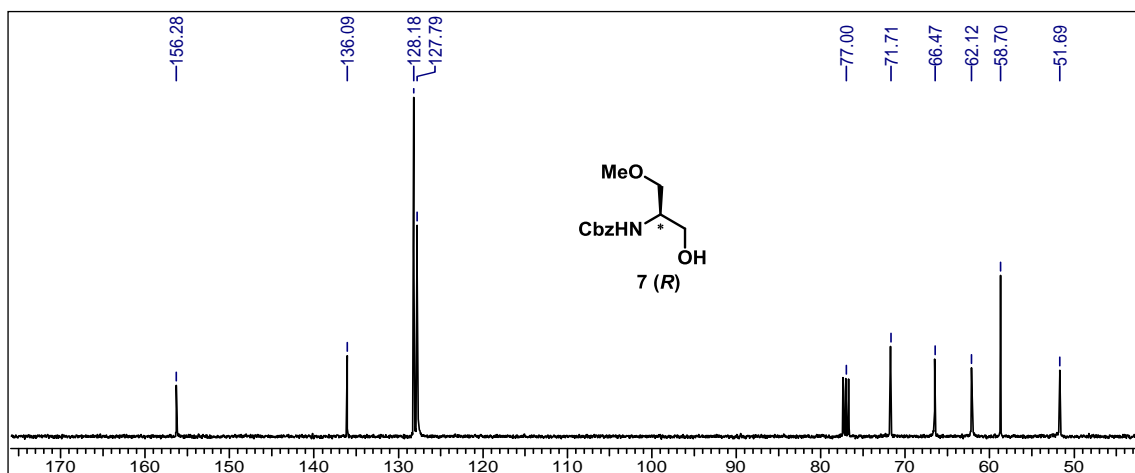
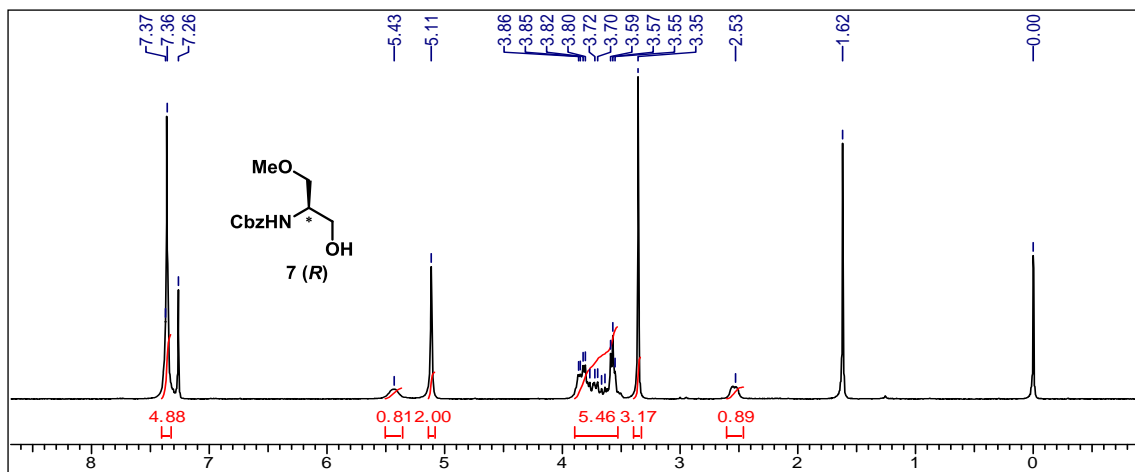
Chiral HPLC chromatograms of enantiomeric tethers from L-serine:

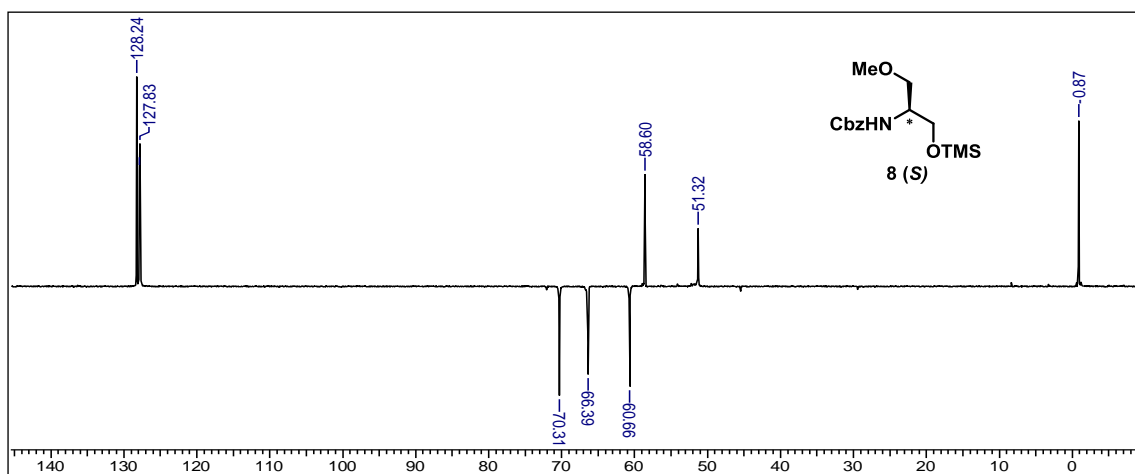
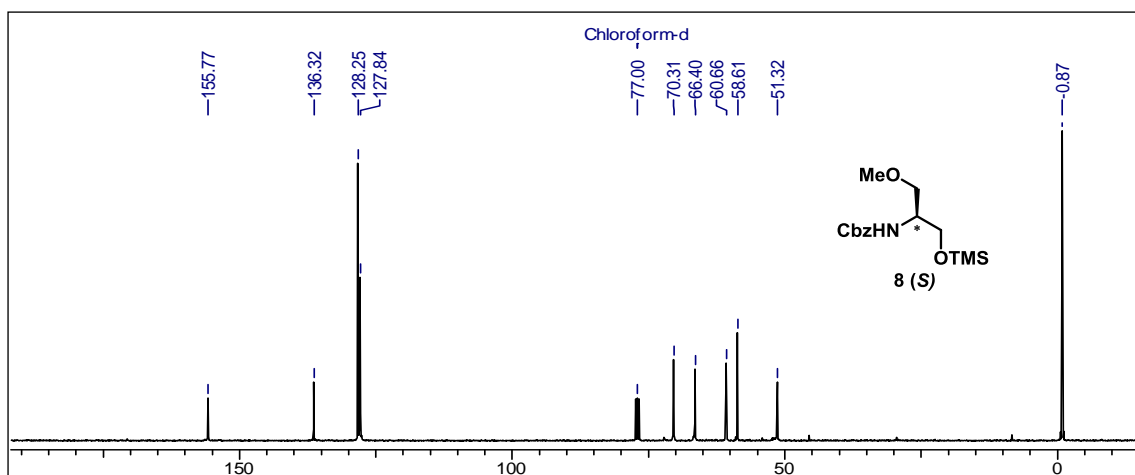
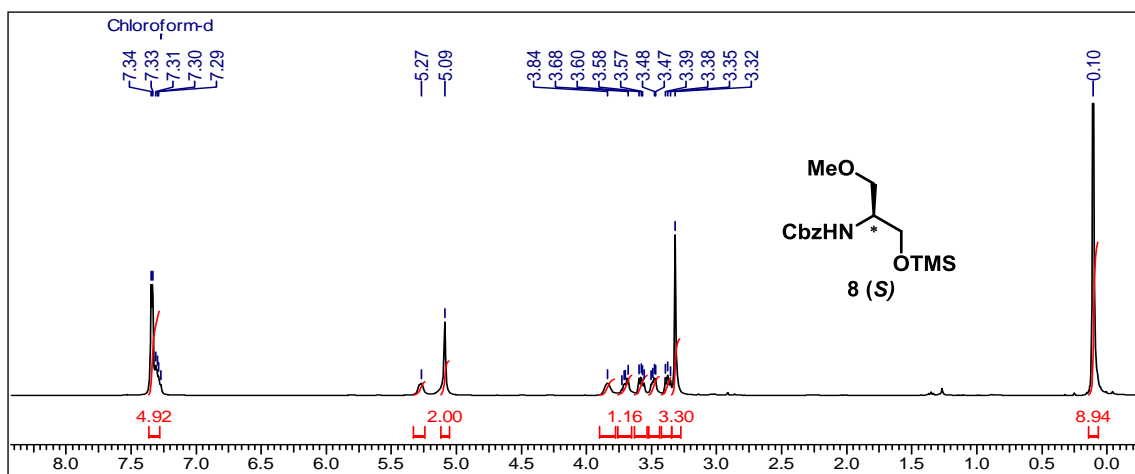
Injection volume of 5 μ L having compound concentration 1mg/ 2mL were run on chiral Kromasil 5-AmyCoat (250 x4.6mm) column with the flow rate of 0.5 mL/min (570 psi) in isocratic system (isopropyl alcohol:n-hexane in 40:60). Comparison of the chromatograms **A** resulted from the injection of **7(R)**, **B** from **7(S)** and **C** from the 1:1 mixture of **7(R)** and **7(S)**, is evidence for the desymmetrization-like concept used for the synthesis of enantiomeric tethers from L-serine.

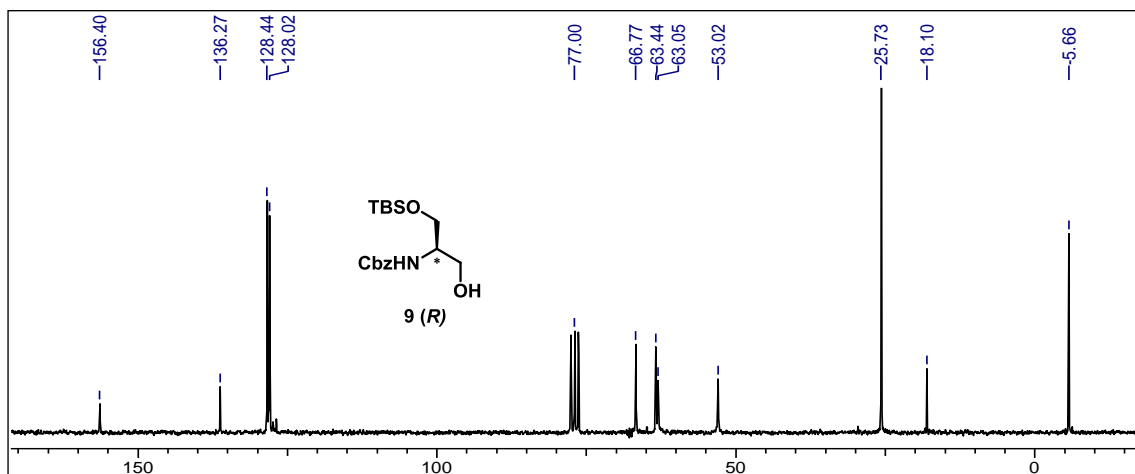
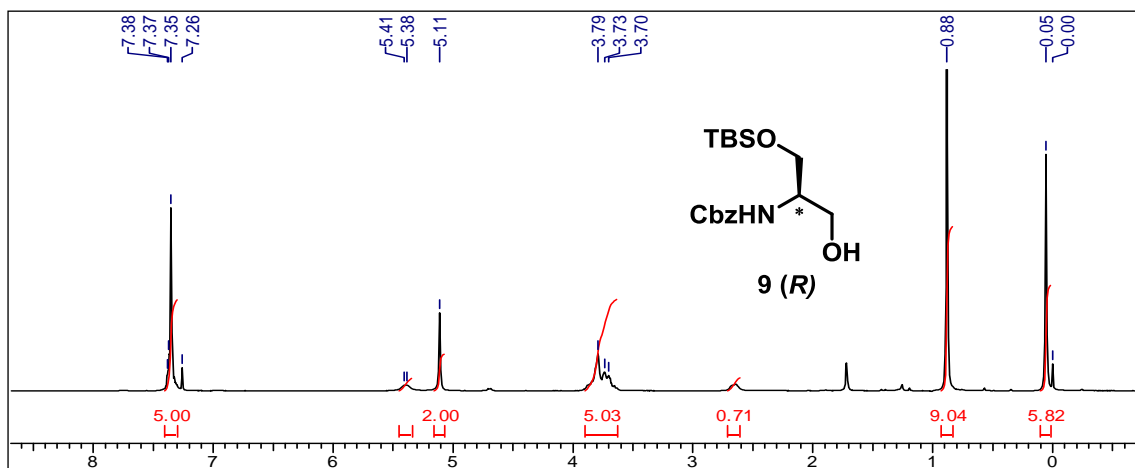
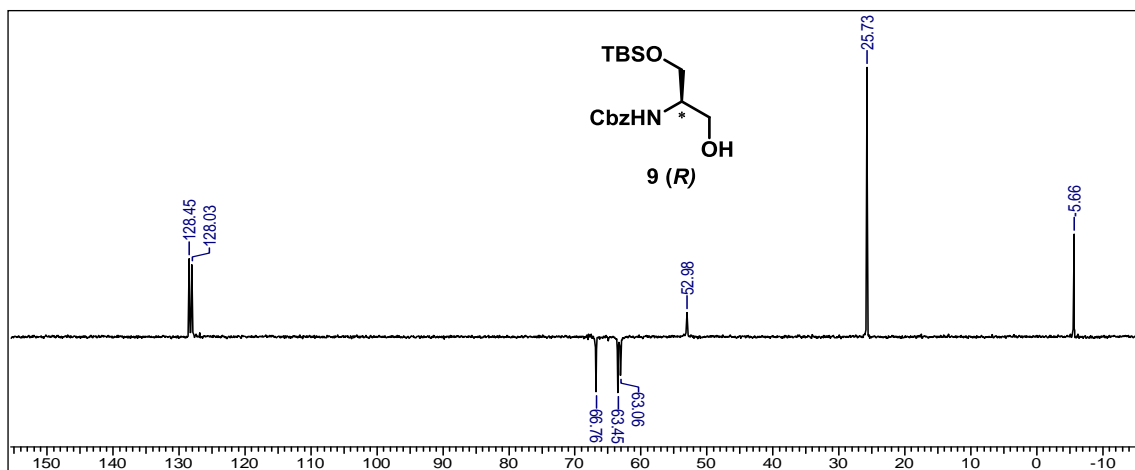


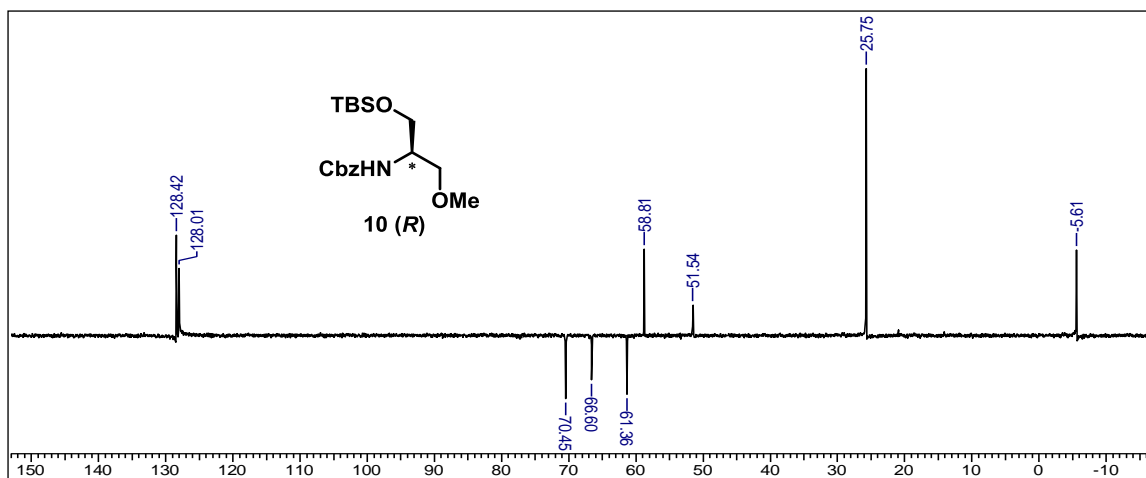
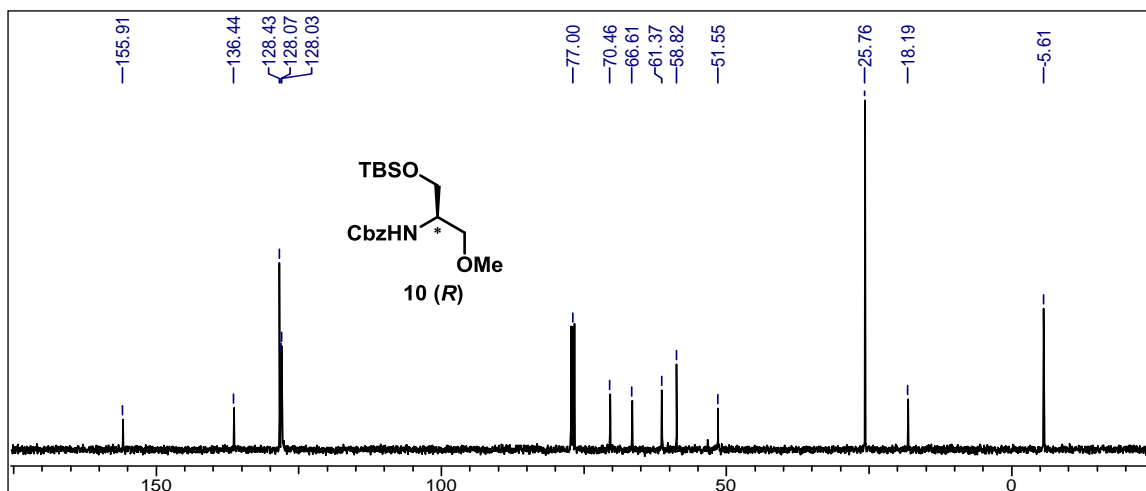
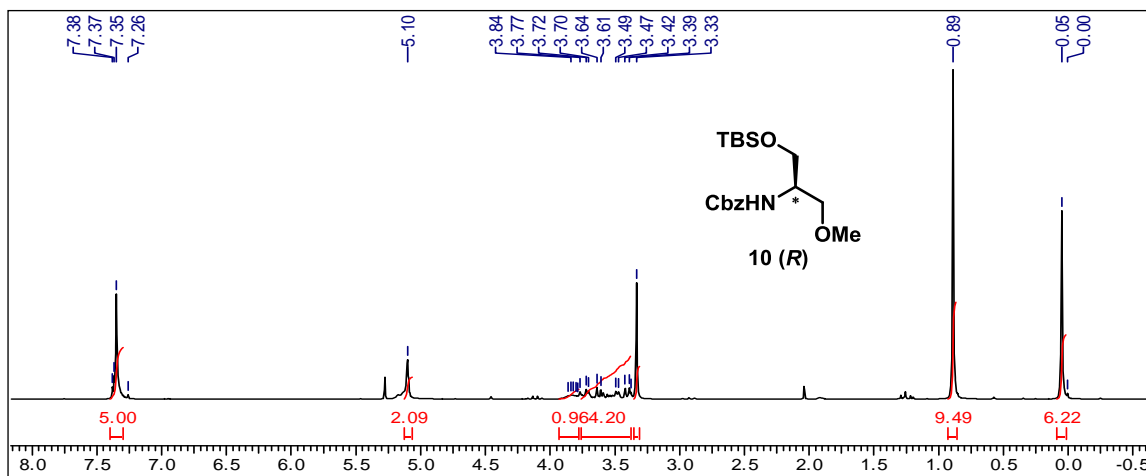
Chapter 2

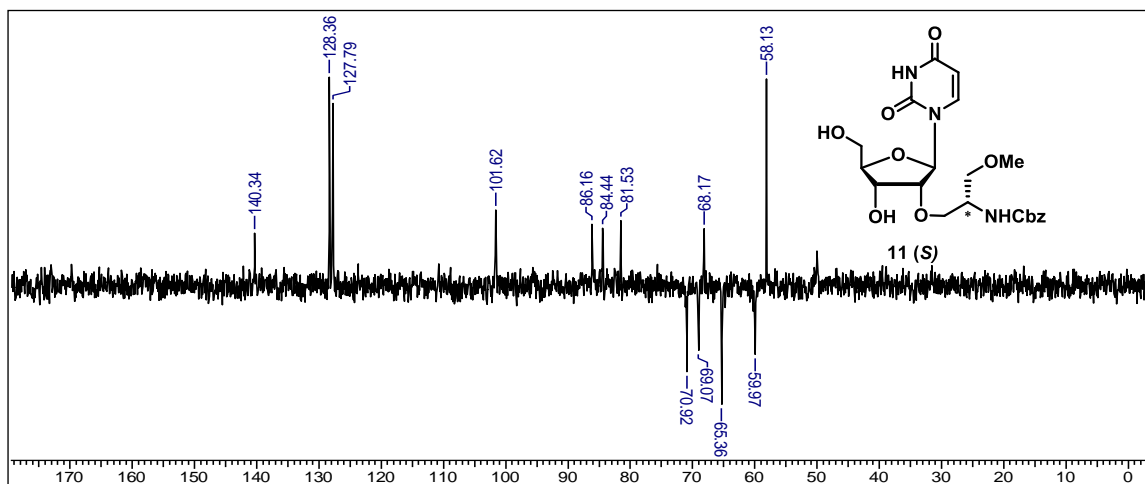
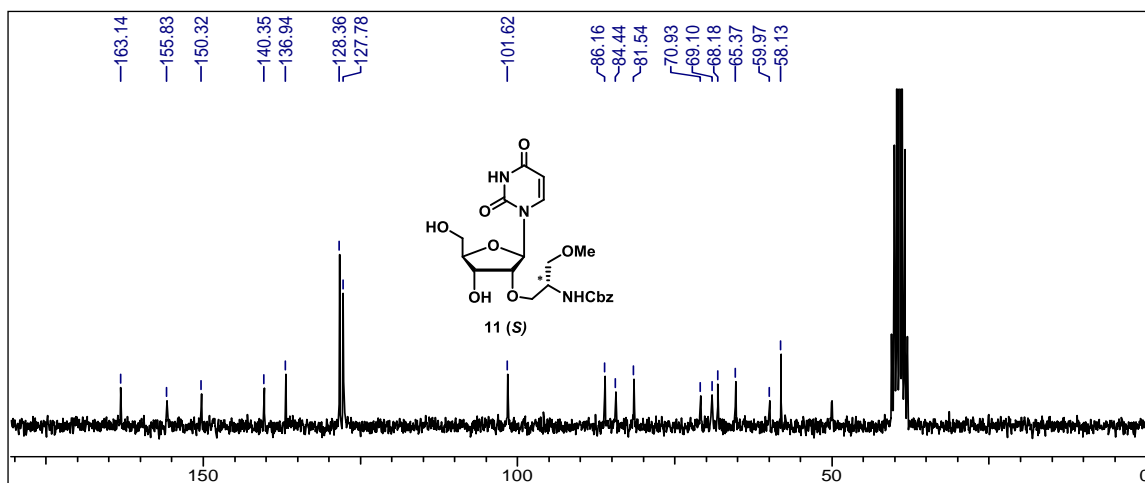
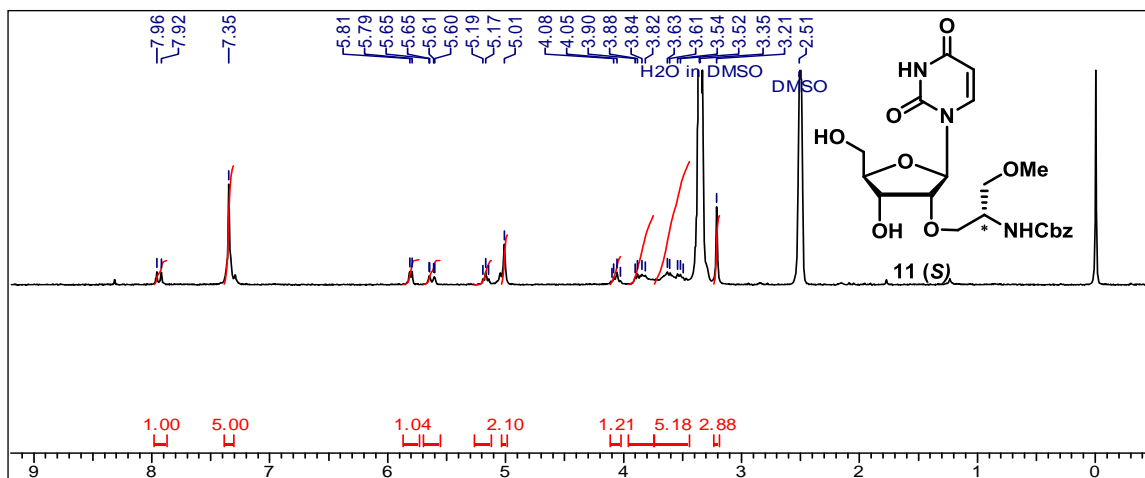
NMR spectral data:

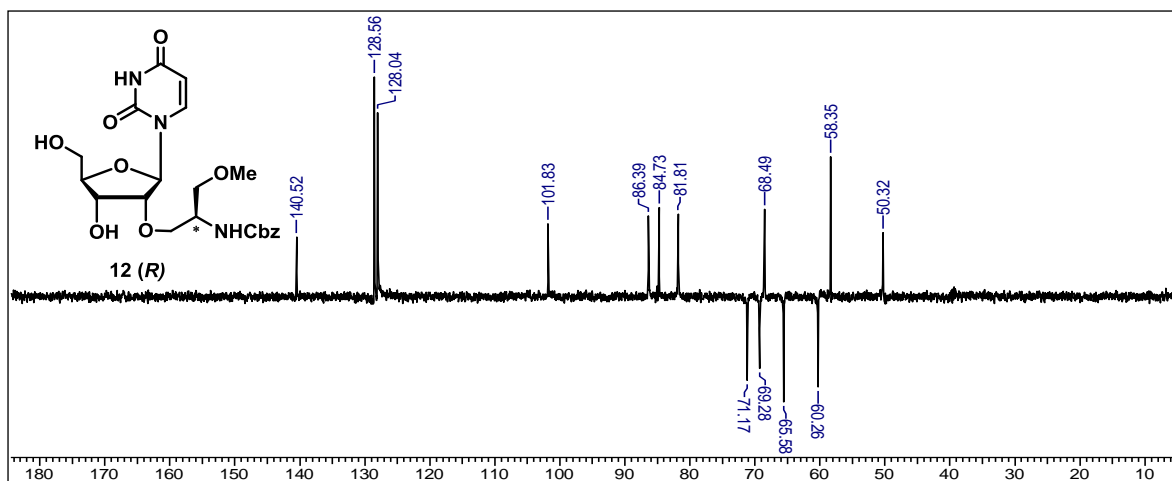
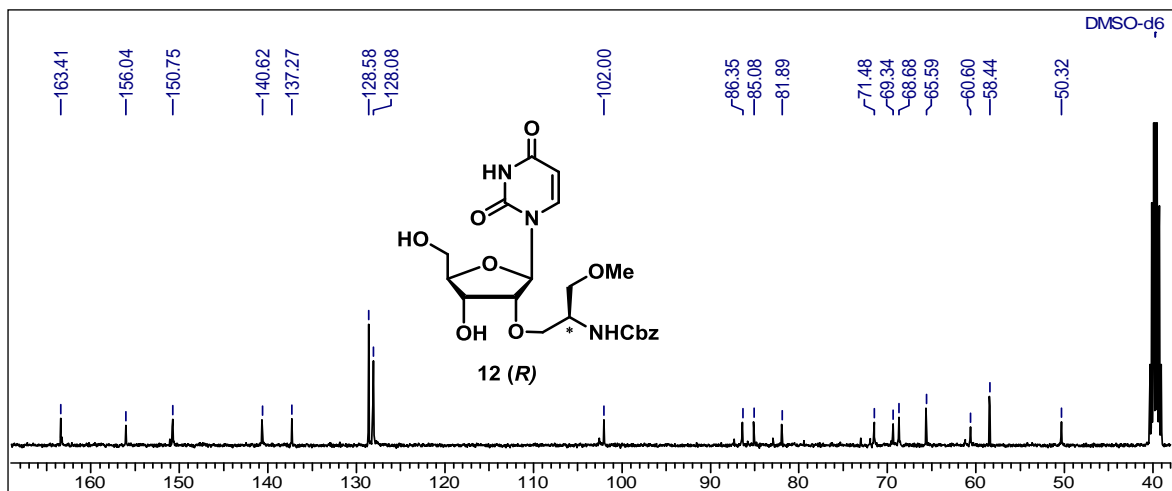
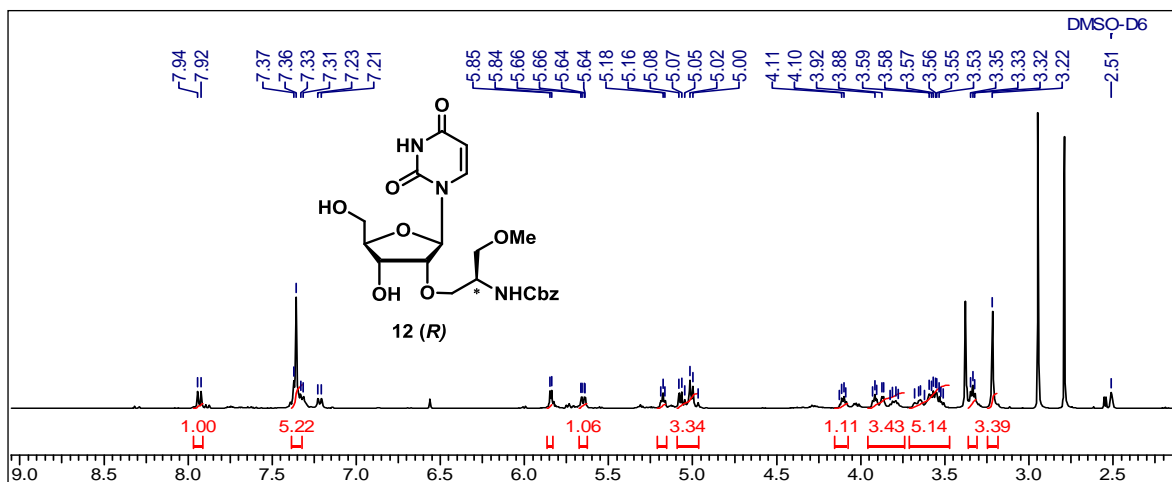


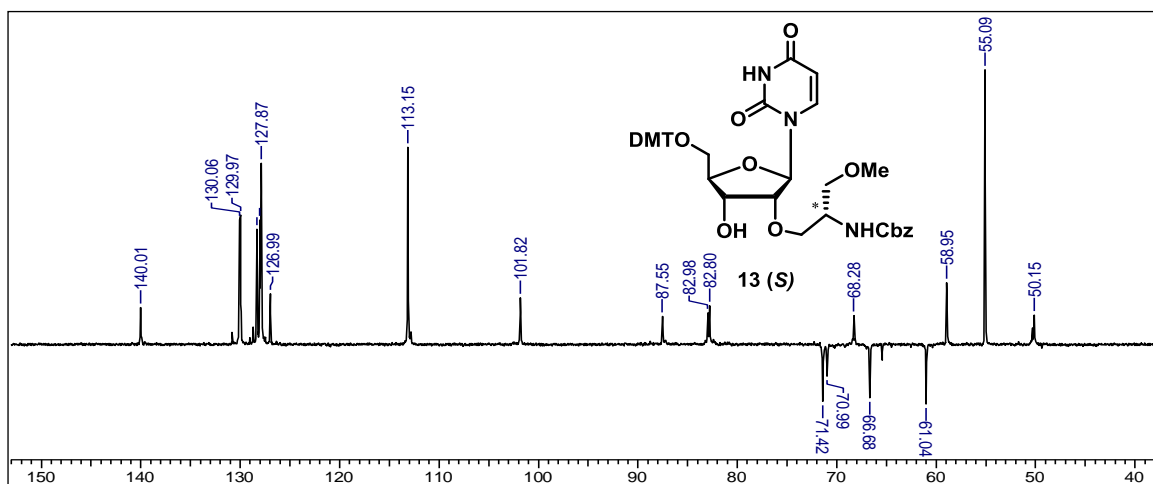
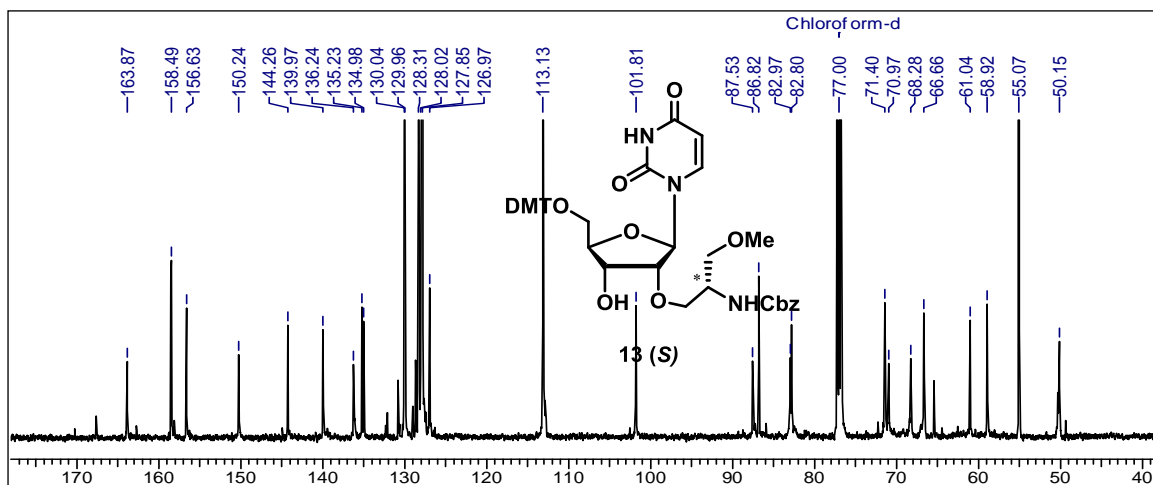
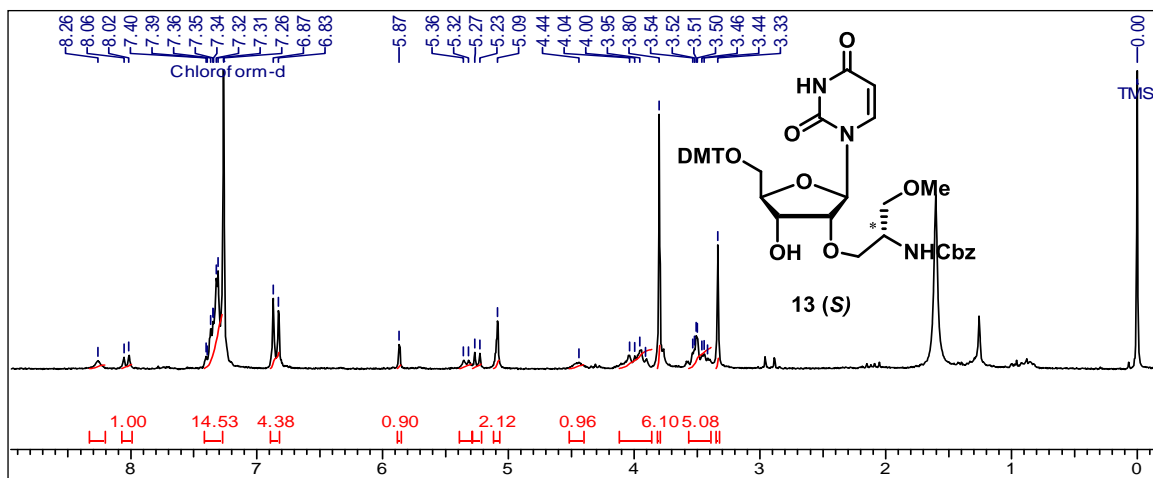




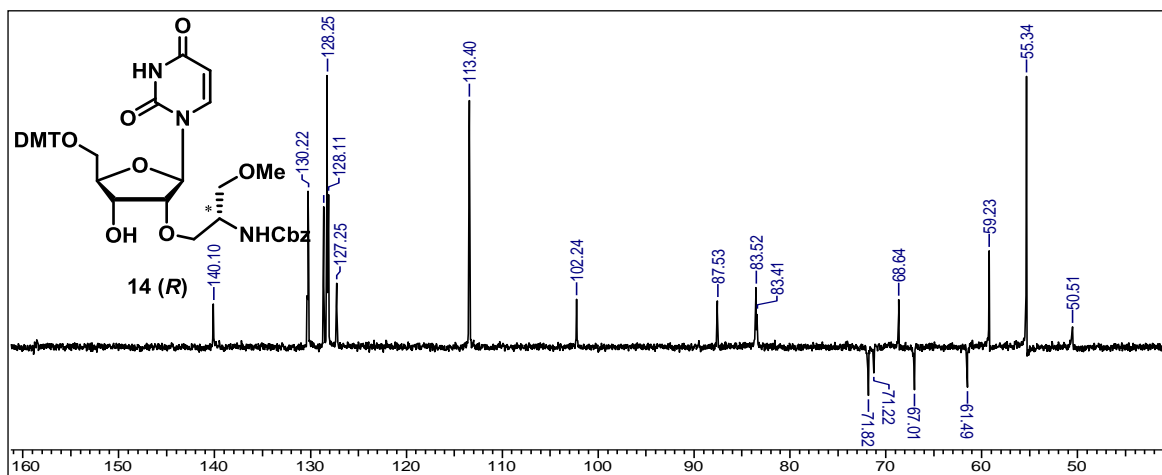
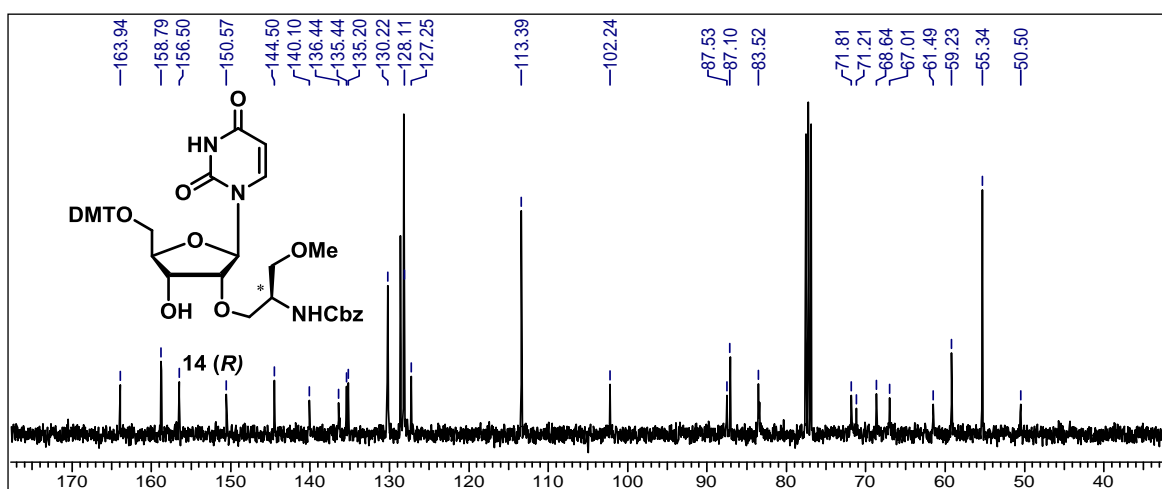
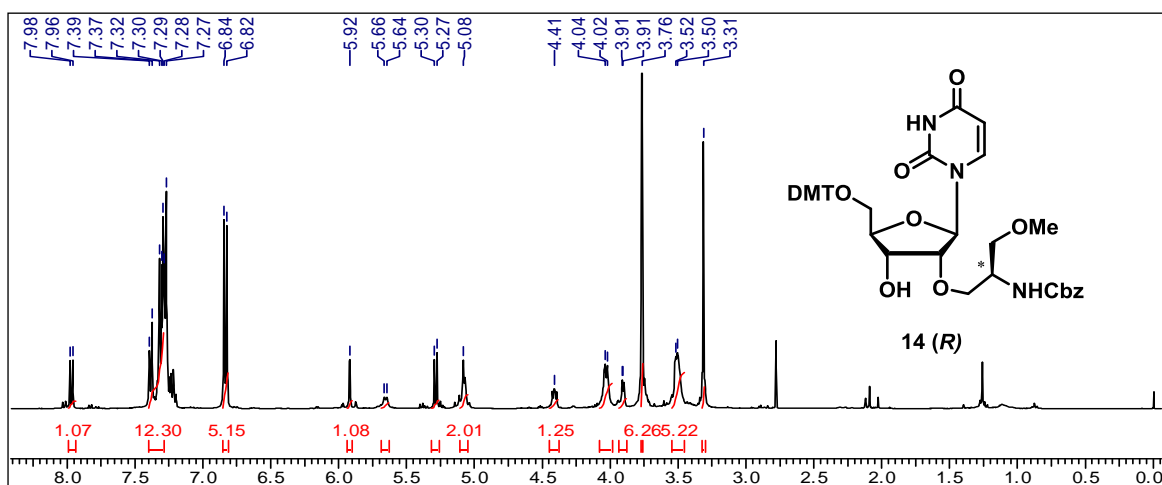




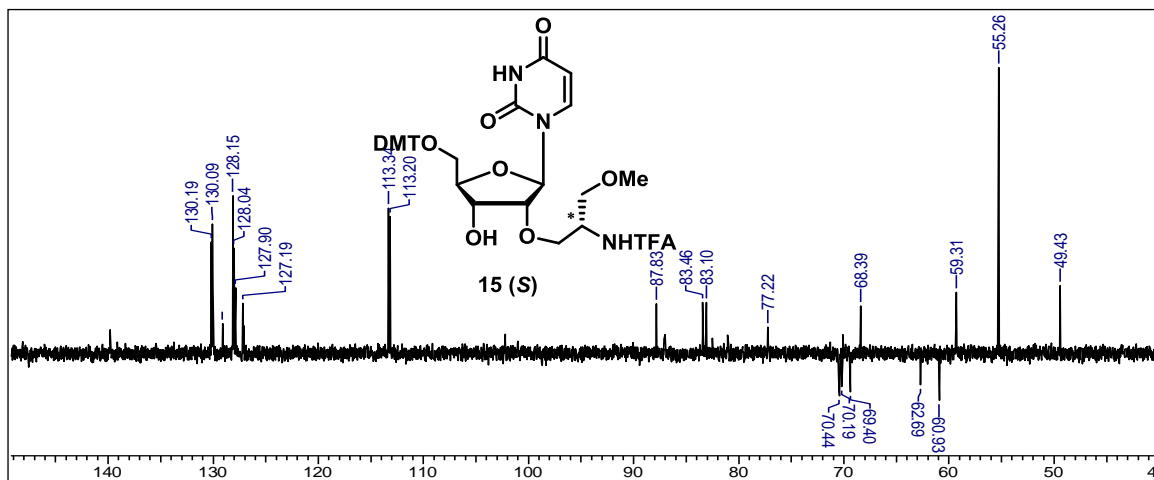
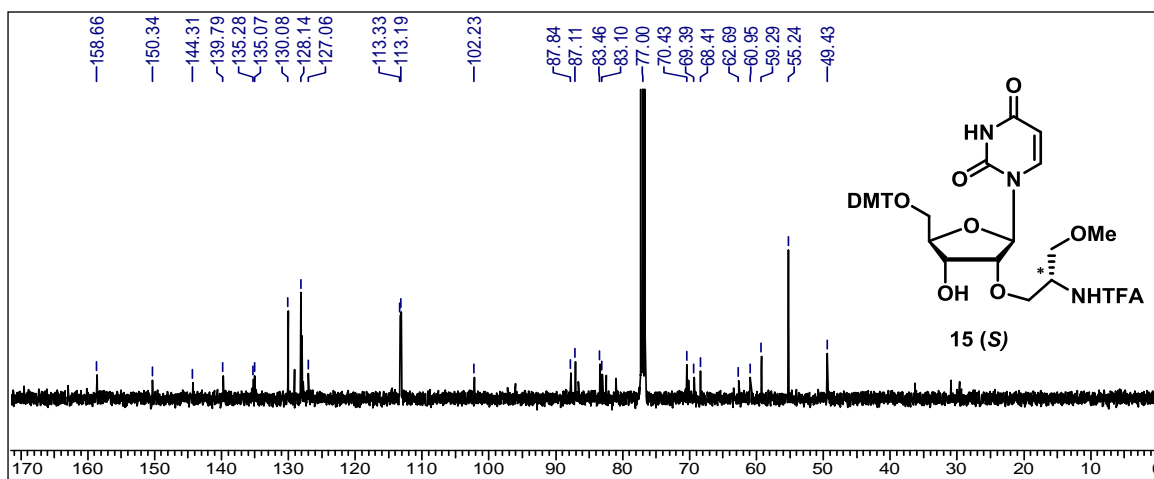
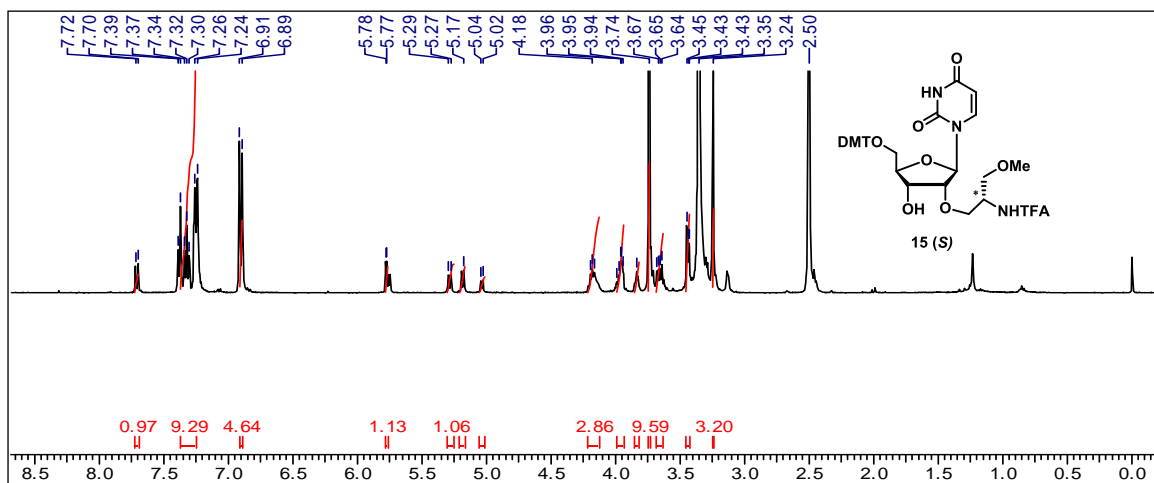




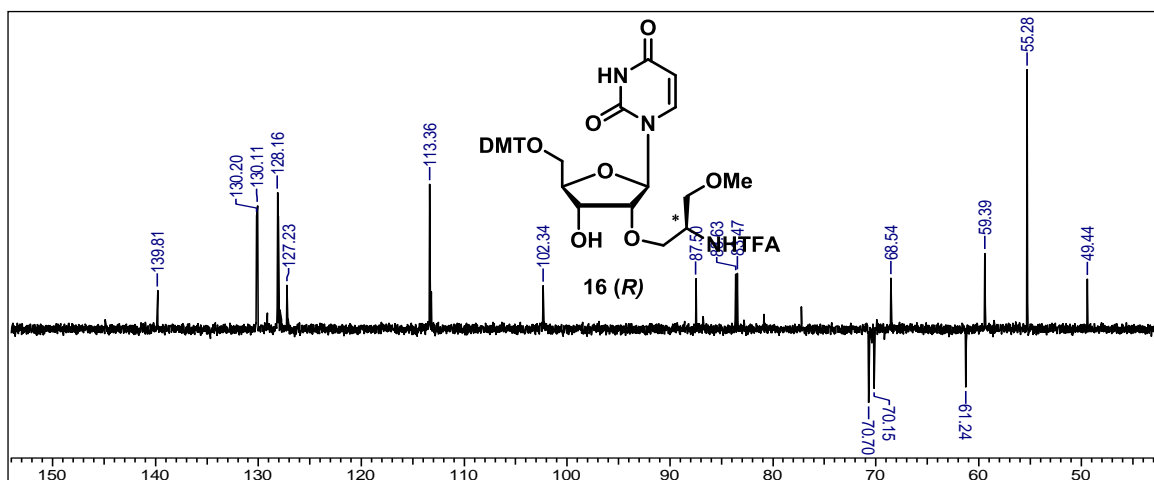
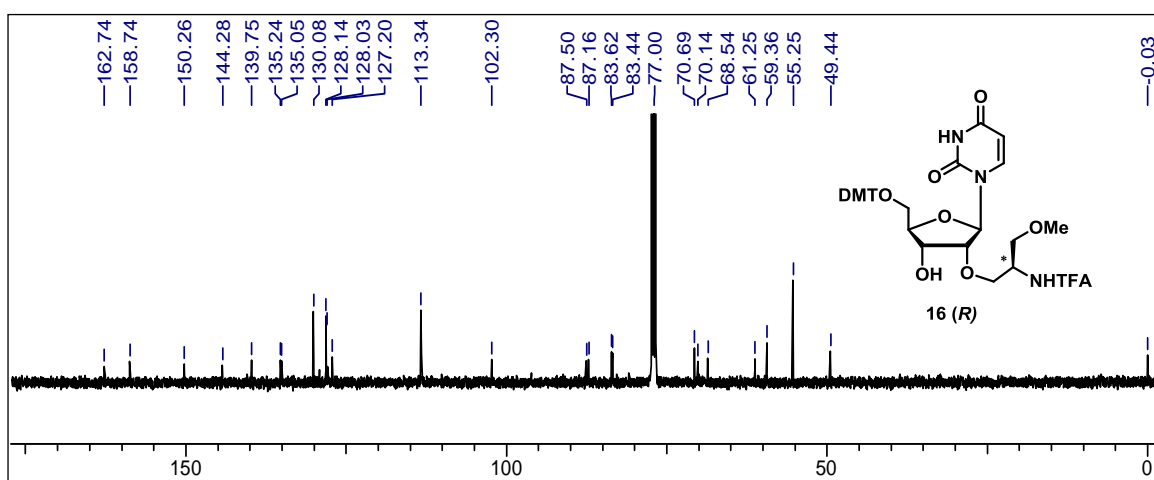
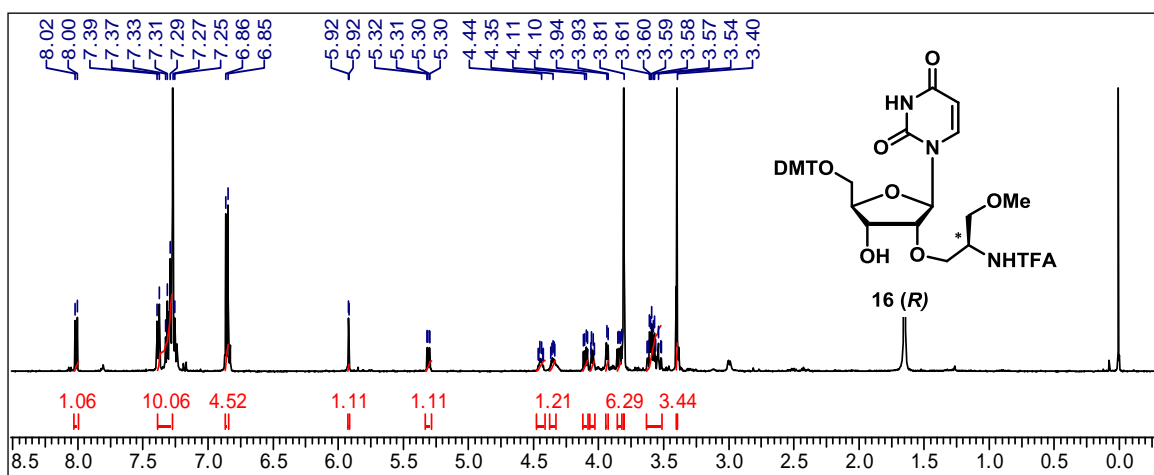
Chapter 2

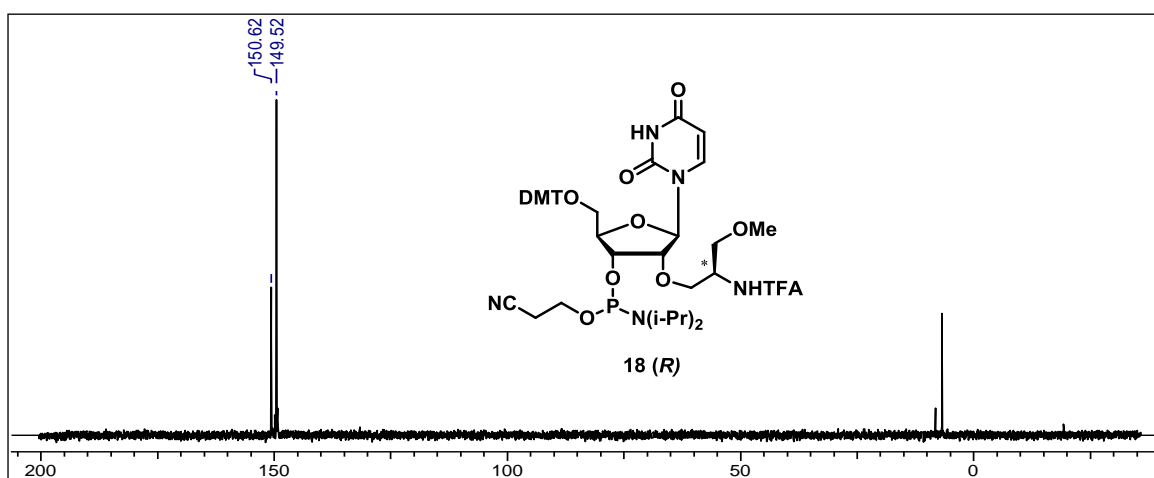
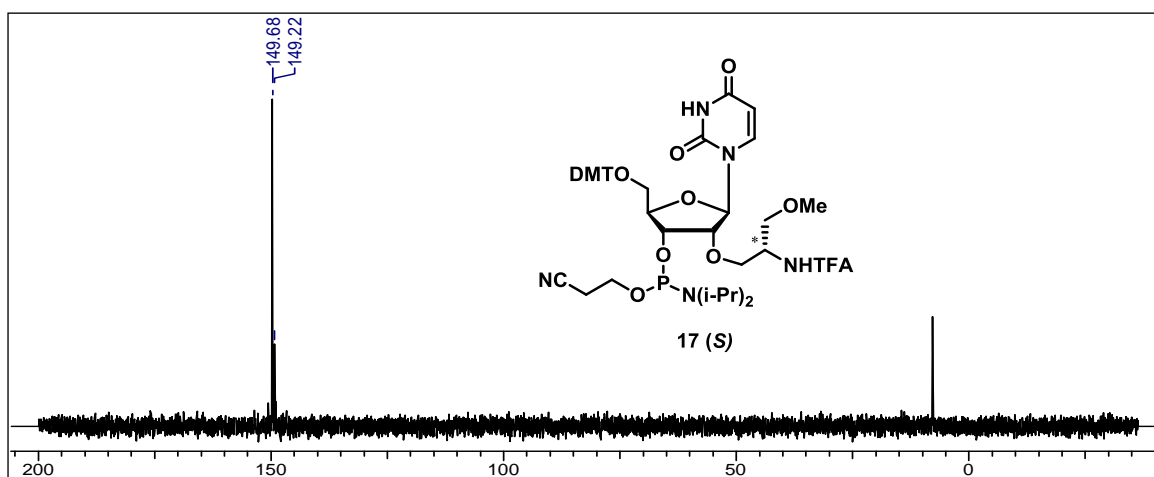


Chapter 2



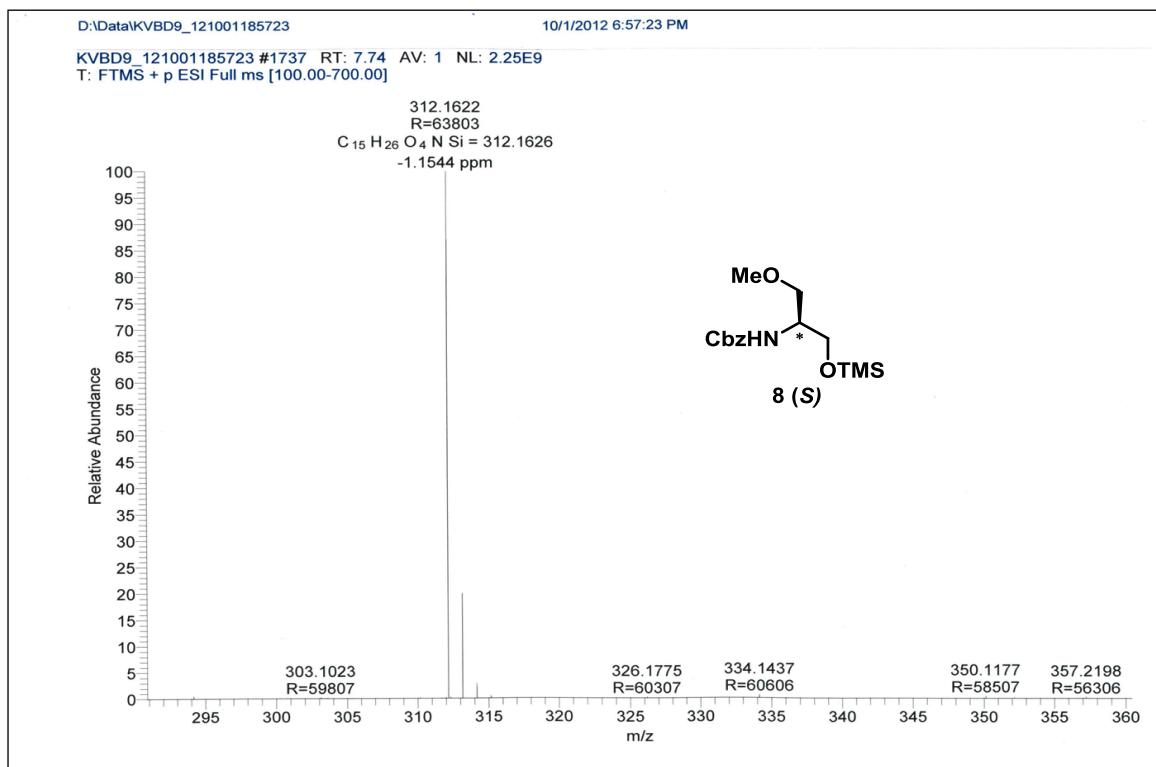
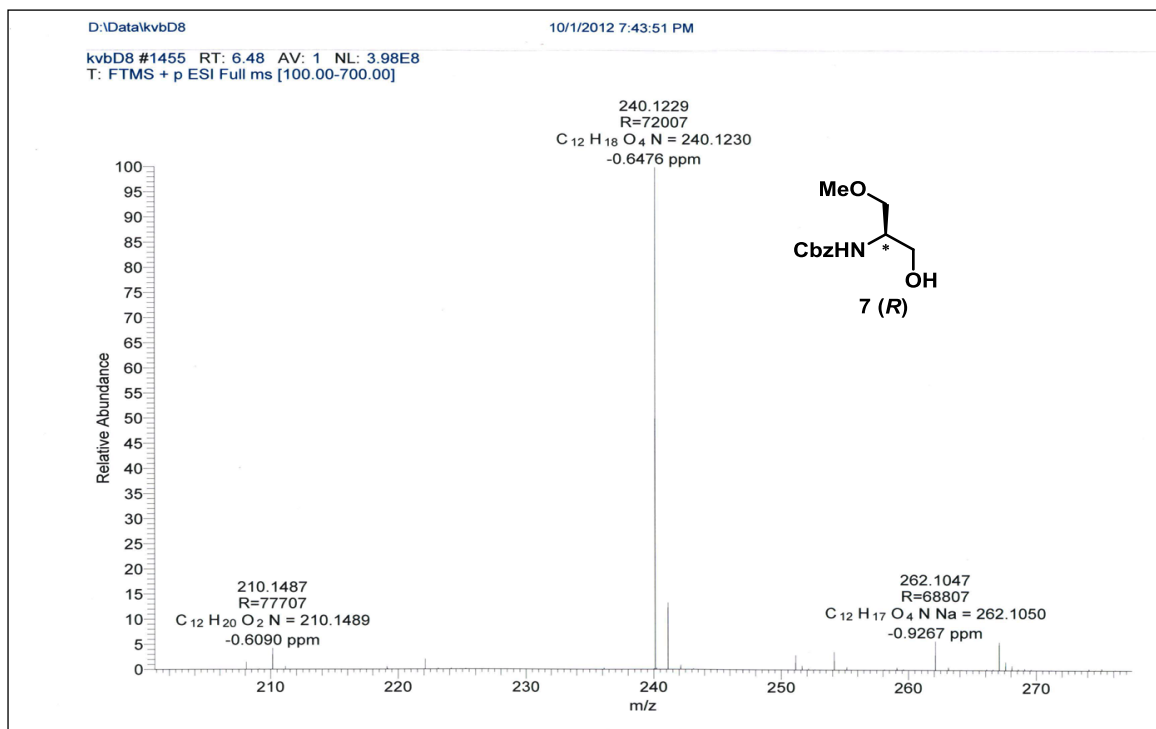
Chapter 2



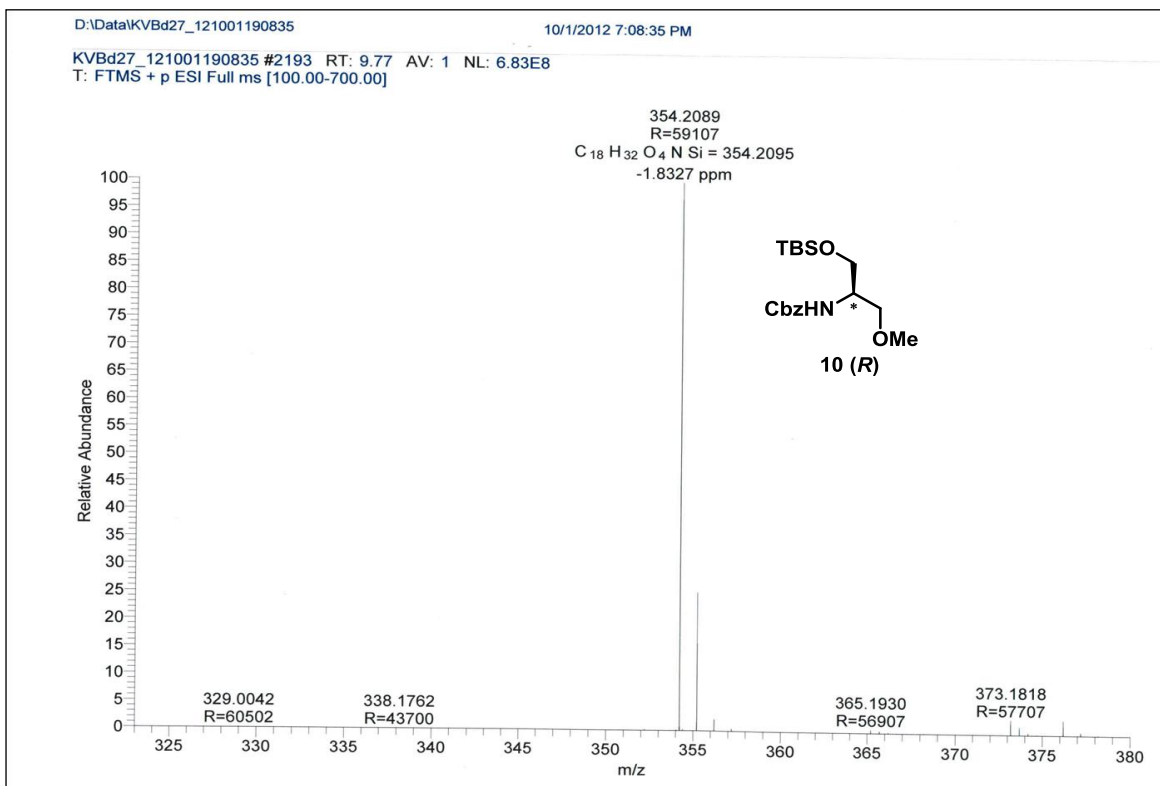
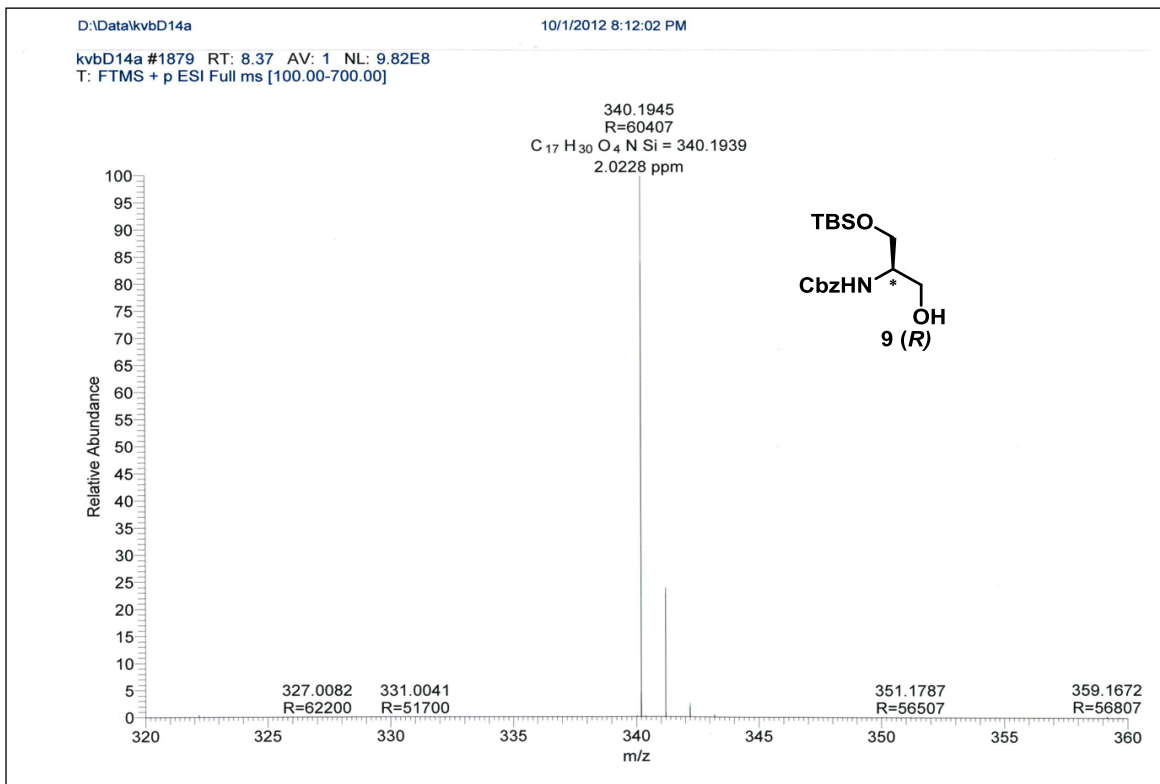


Chapter 2

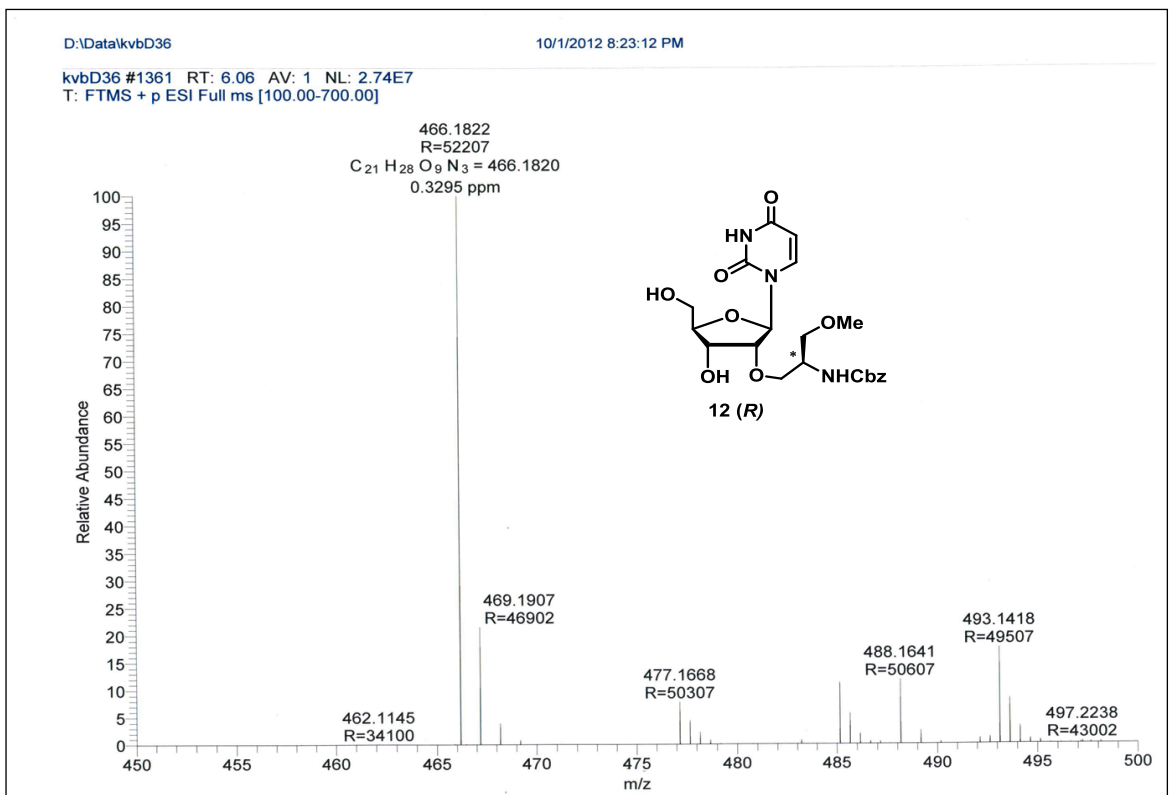
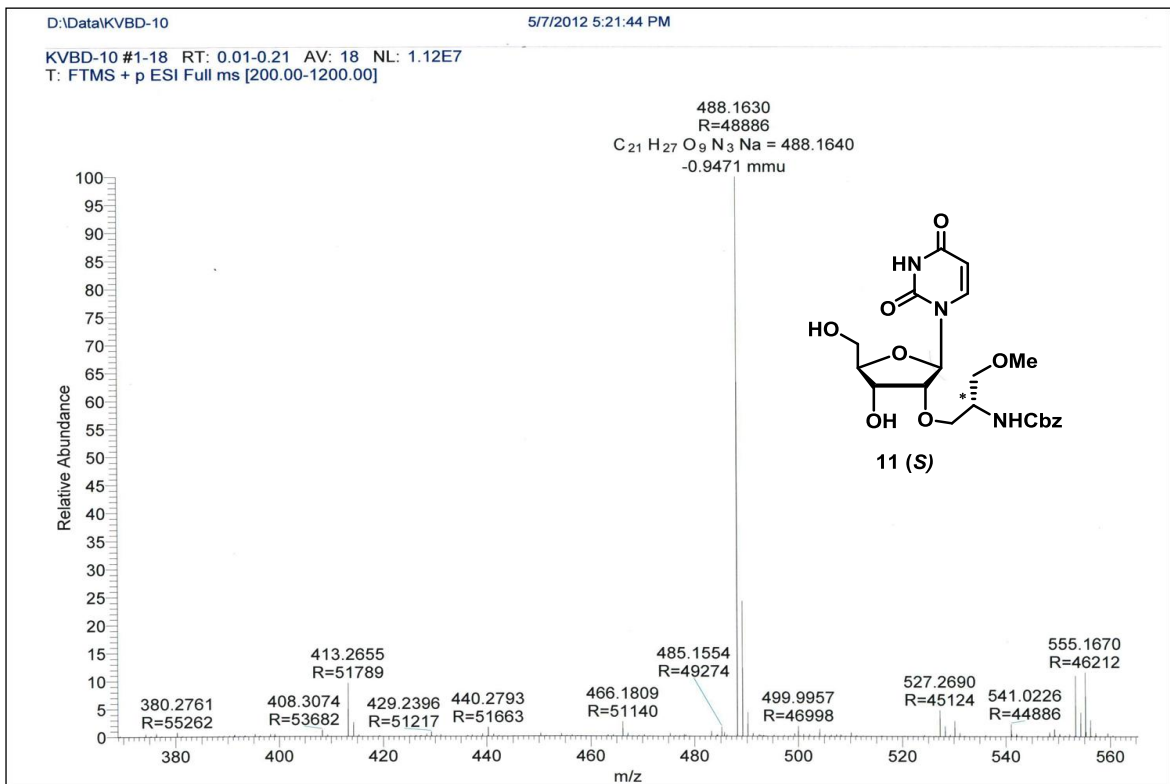
HRMS spectral data:



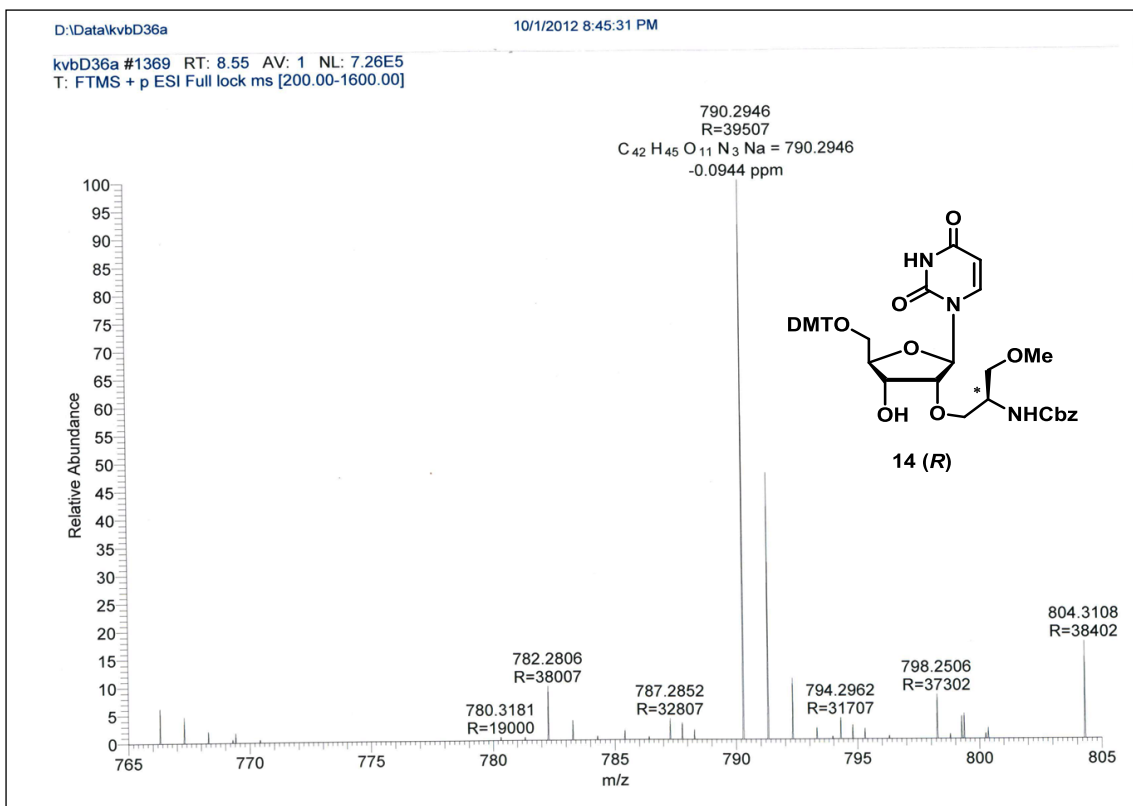
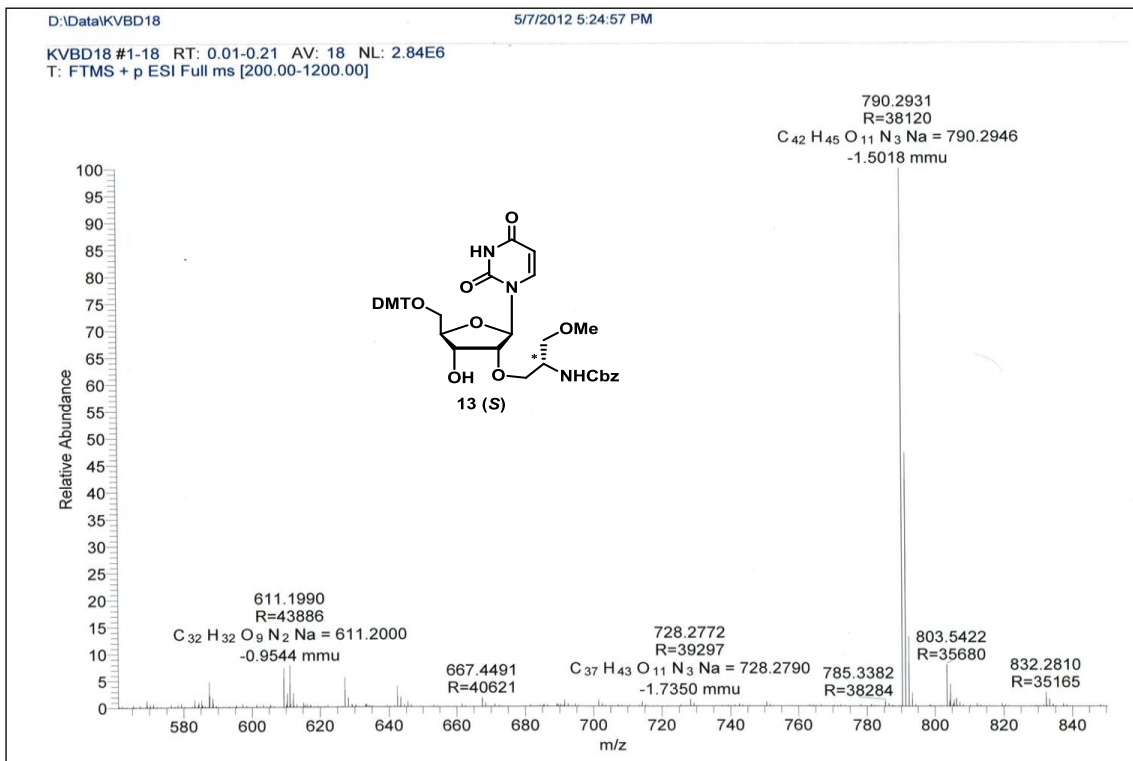
Chapter 2



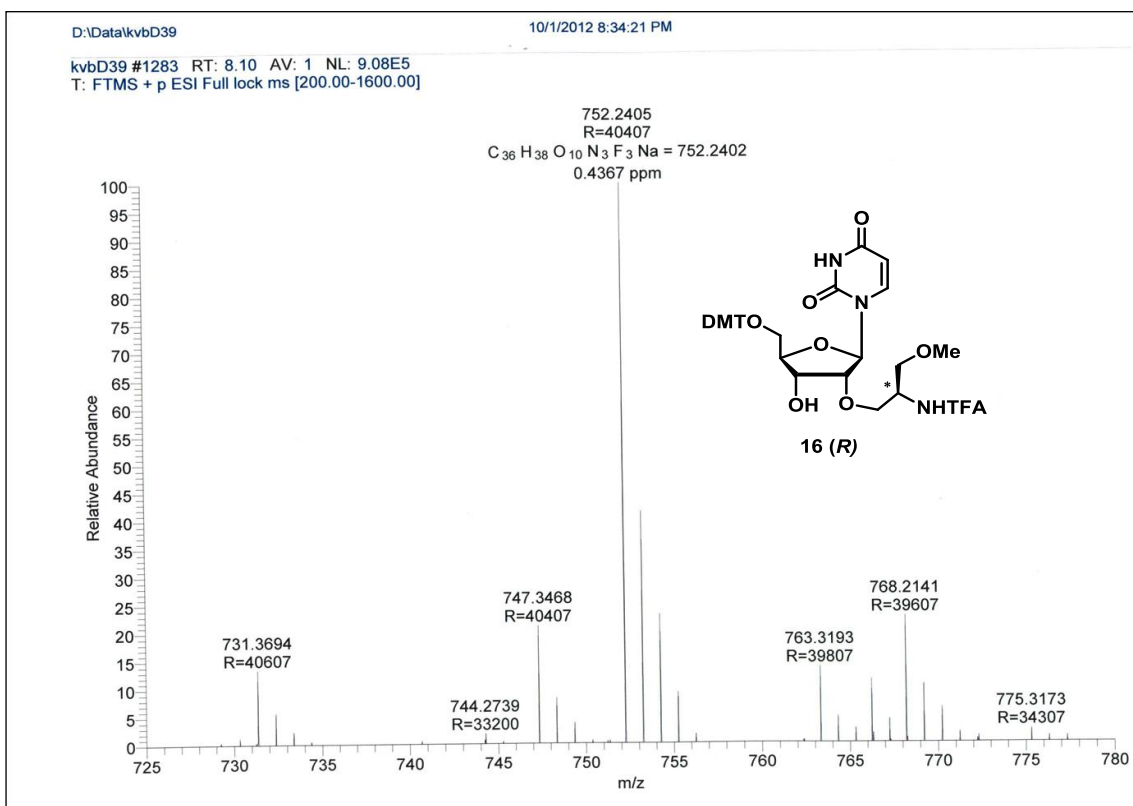
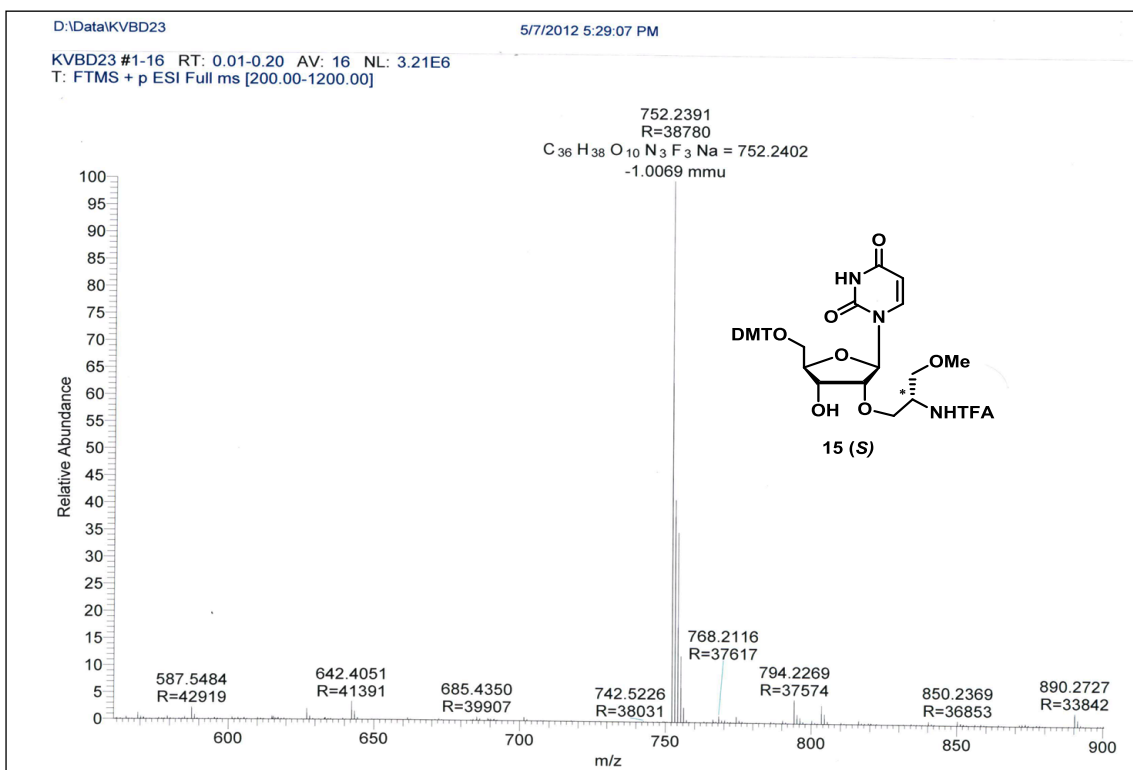
Chapter 2



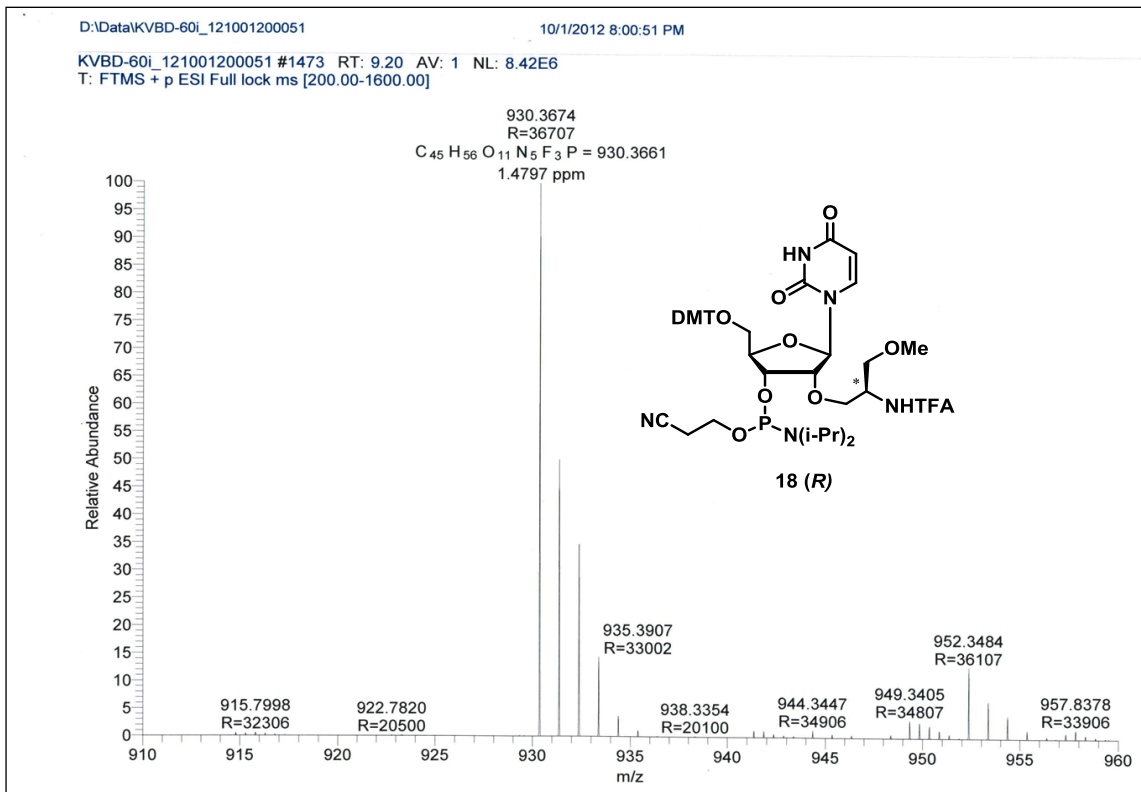
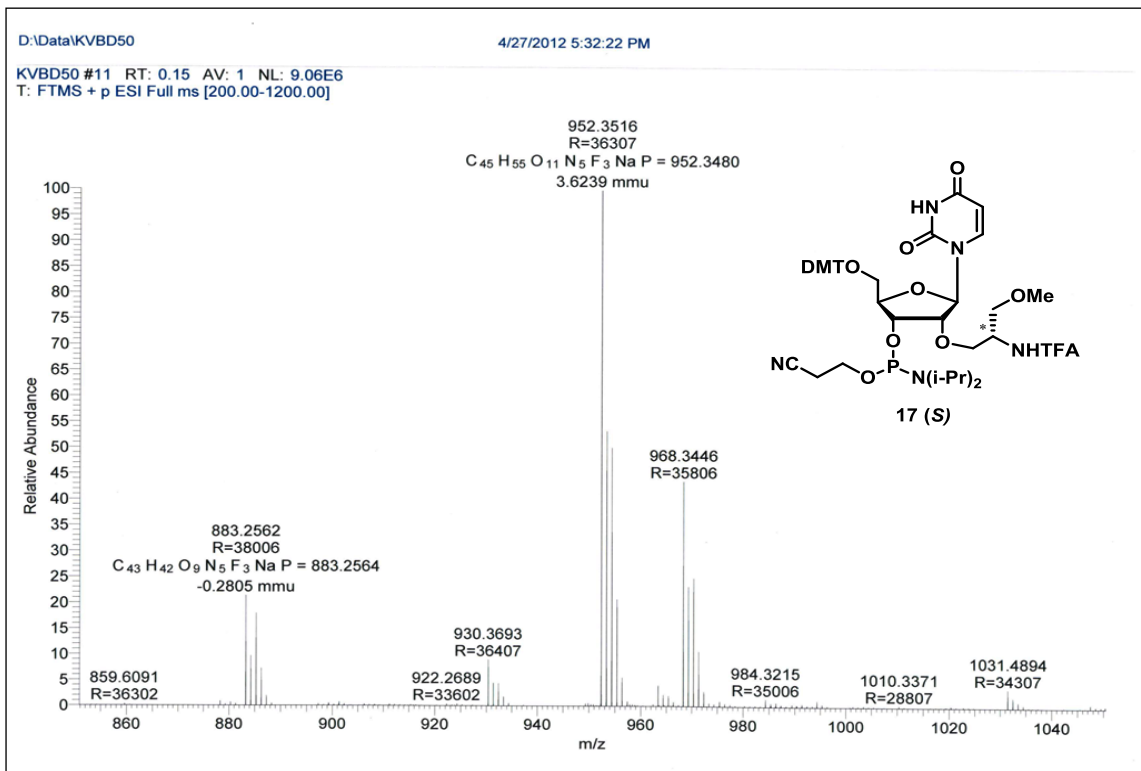
Chapter 2



Chapter 2

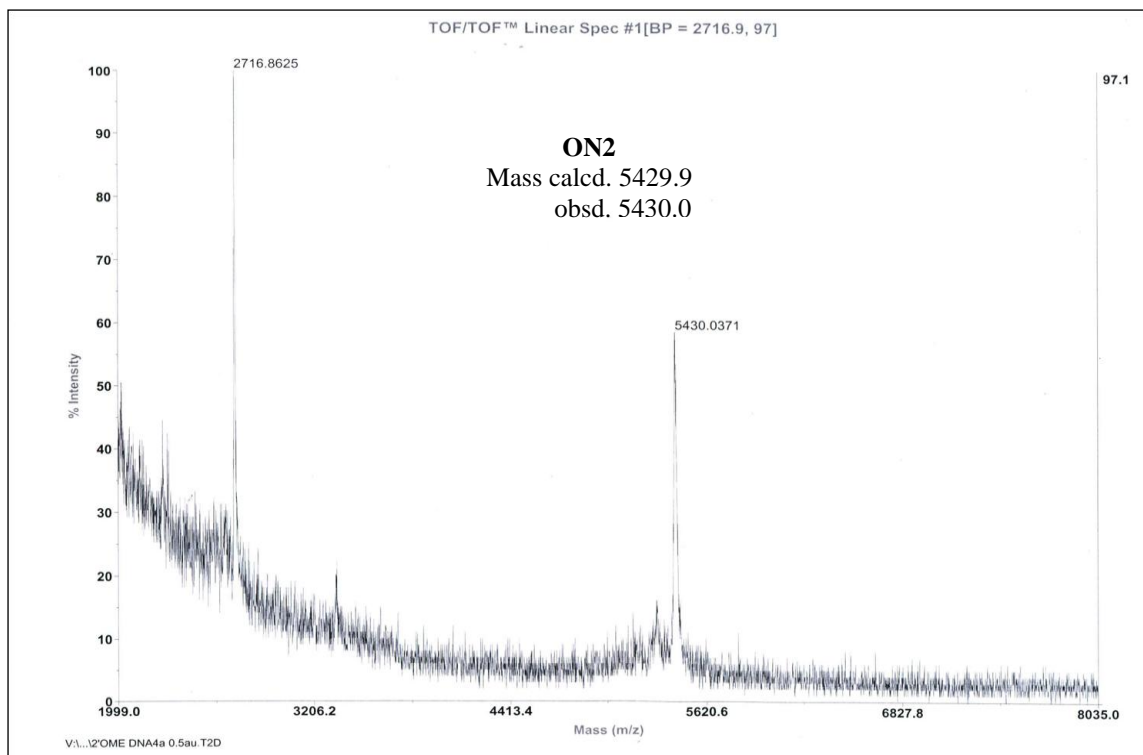
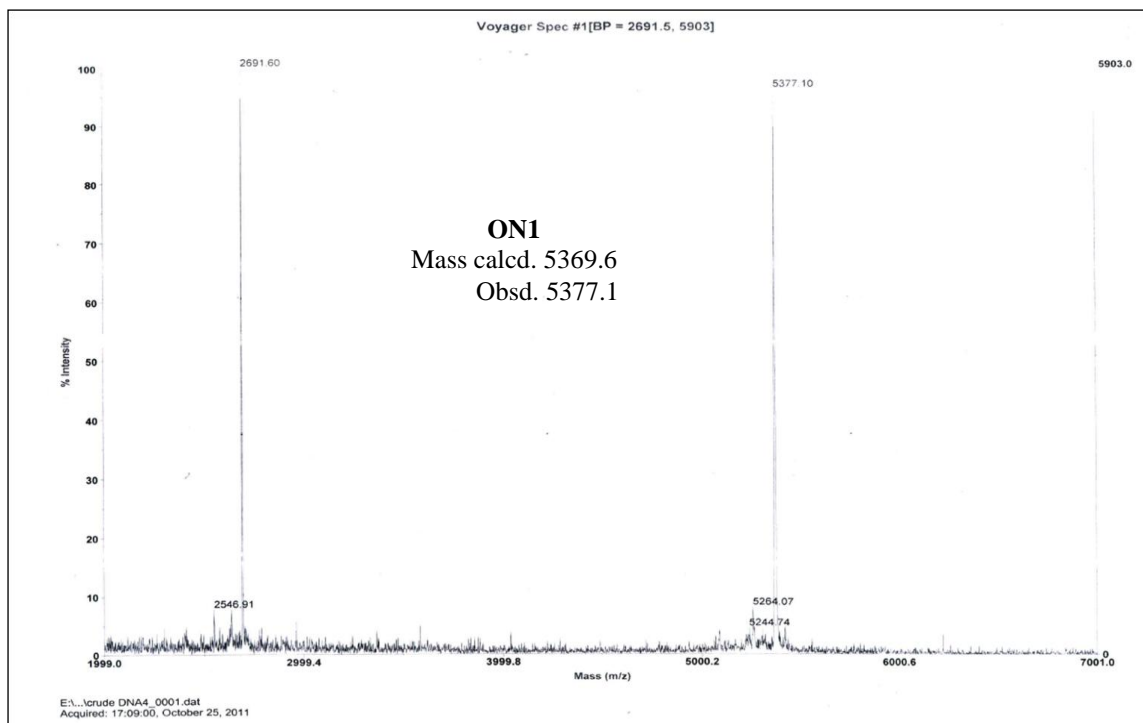


Chapter 2

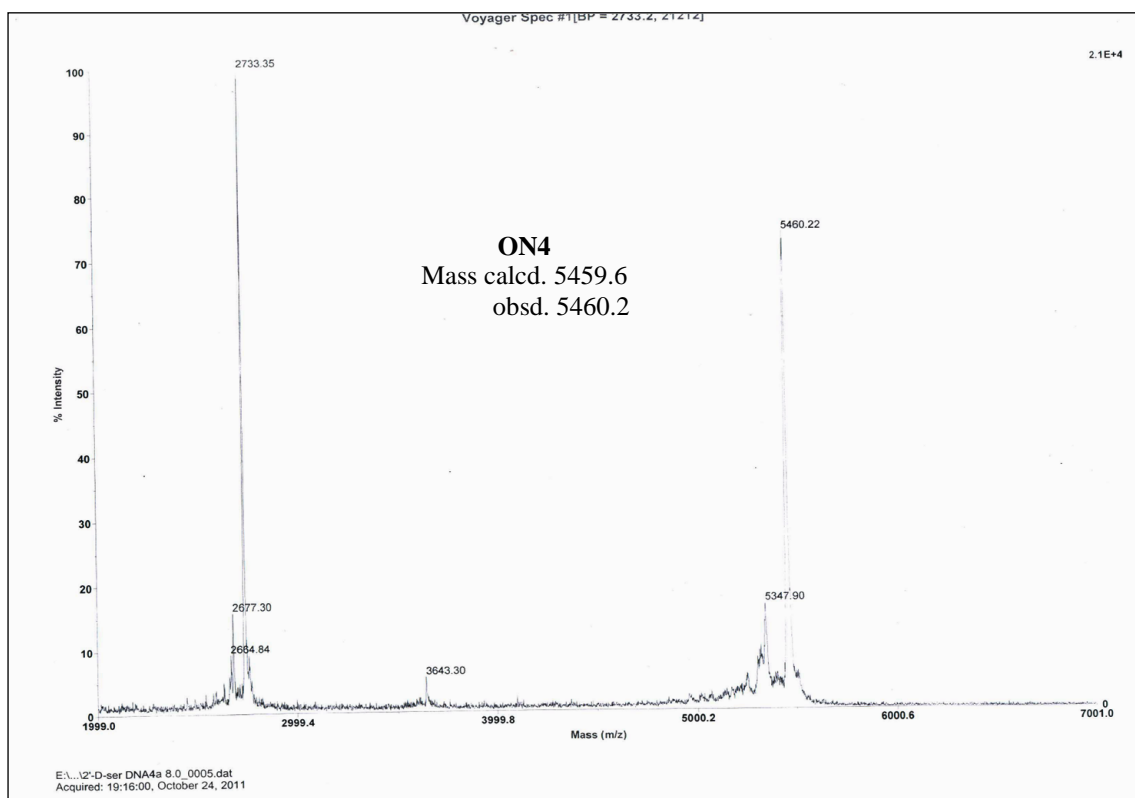
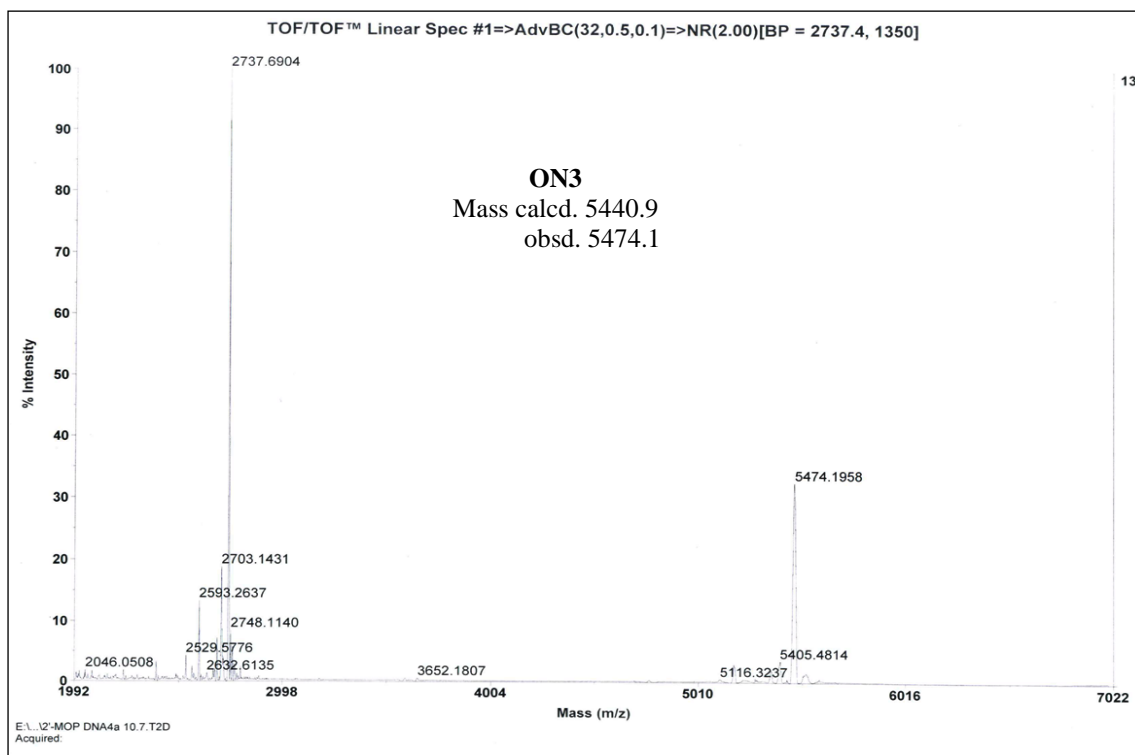


Chapter 2

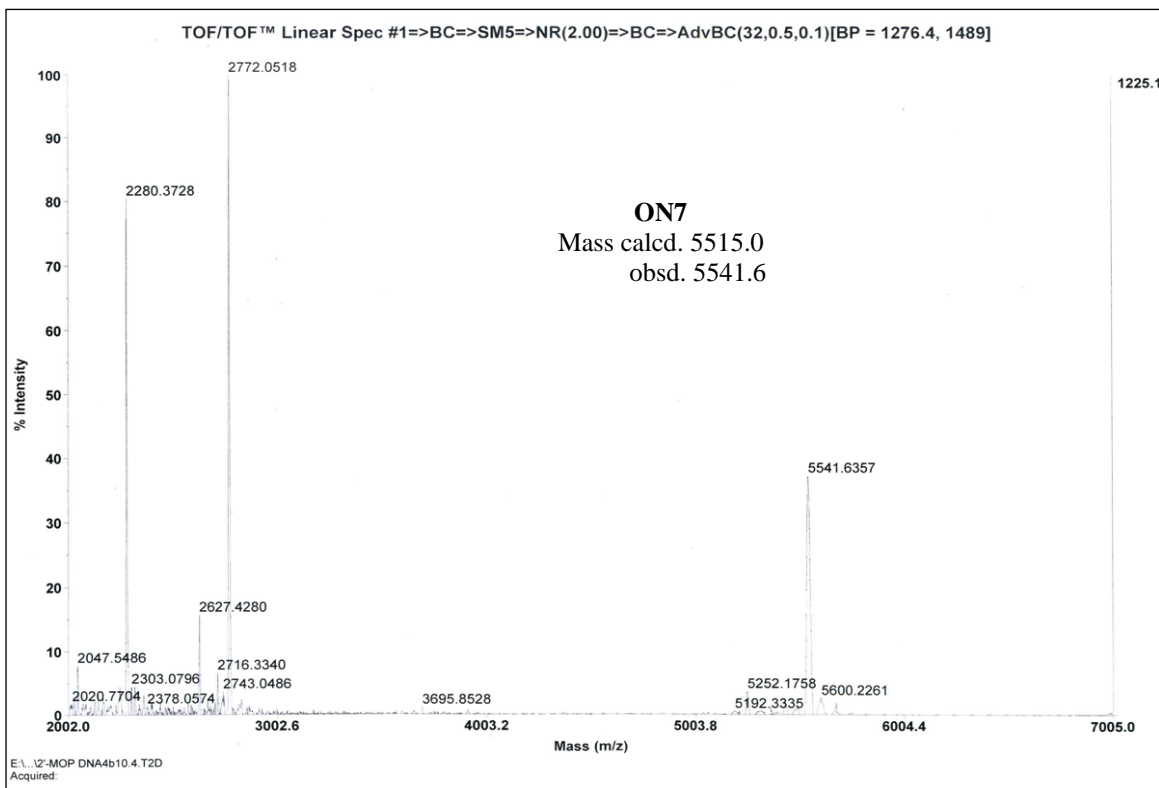
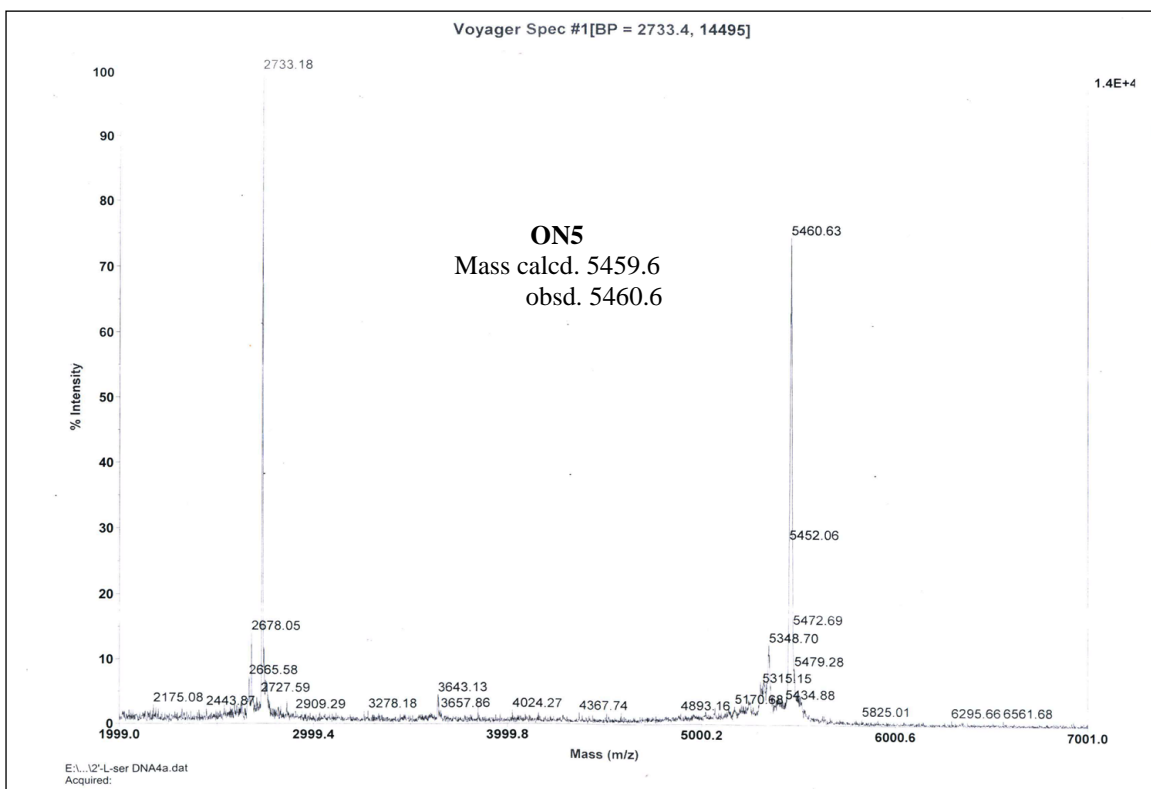
MALDI-TOF spectra of oligonucleotides:



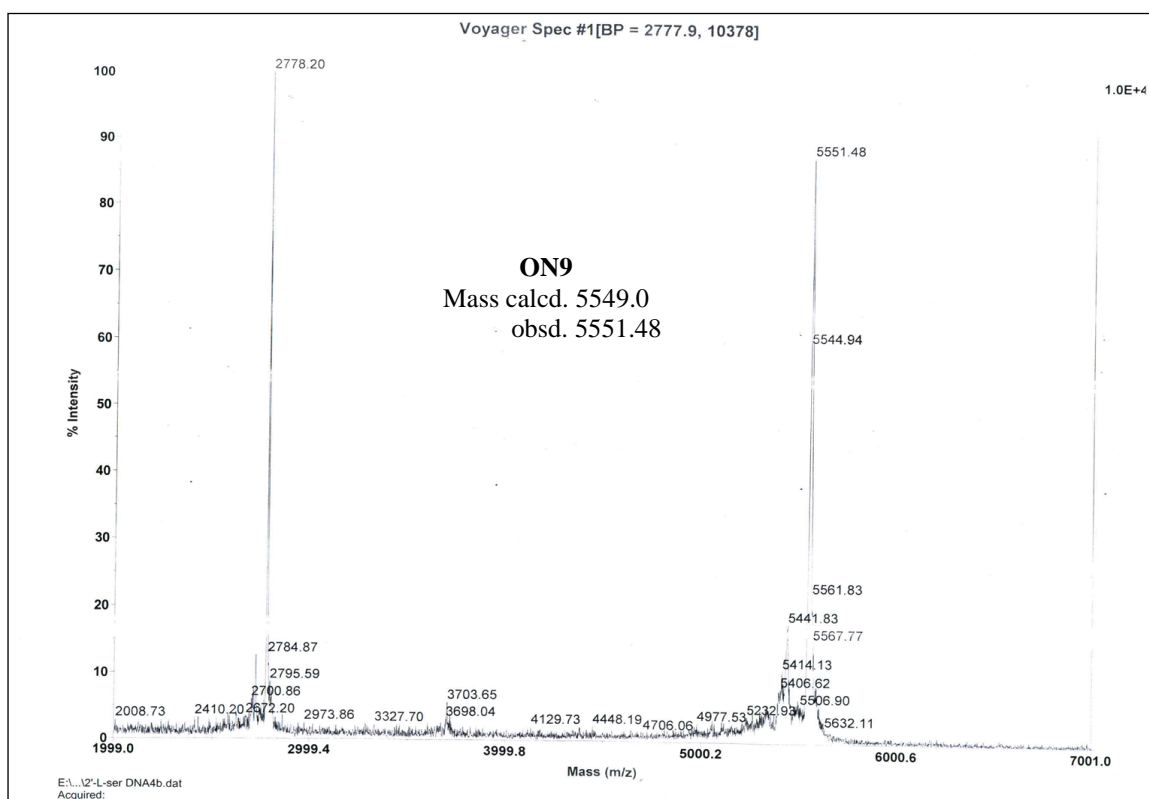
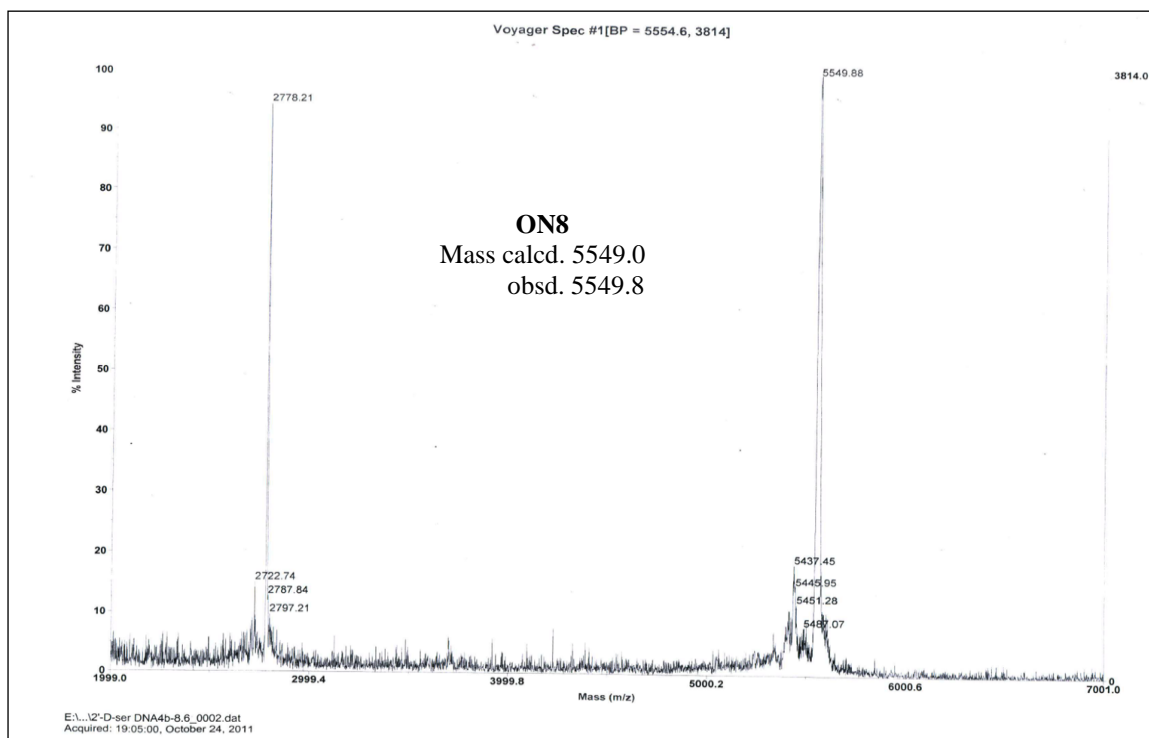
Chapter 2



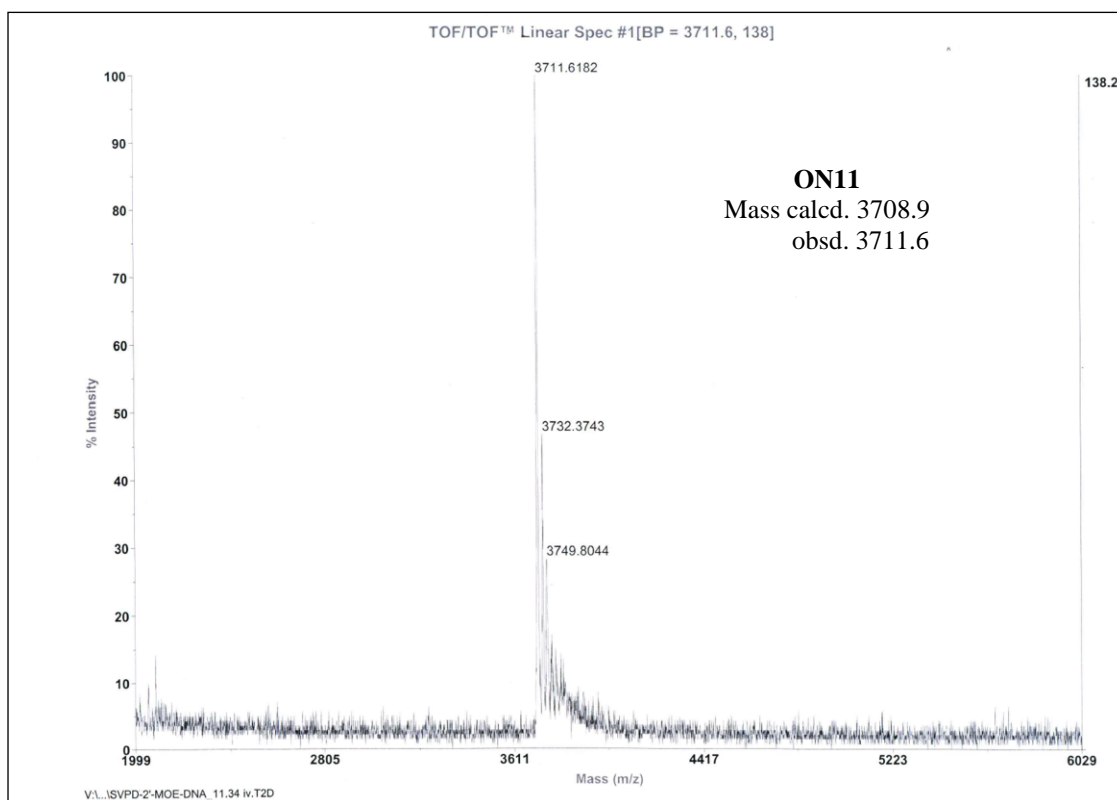
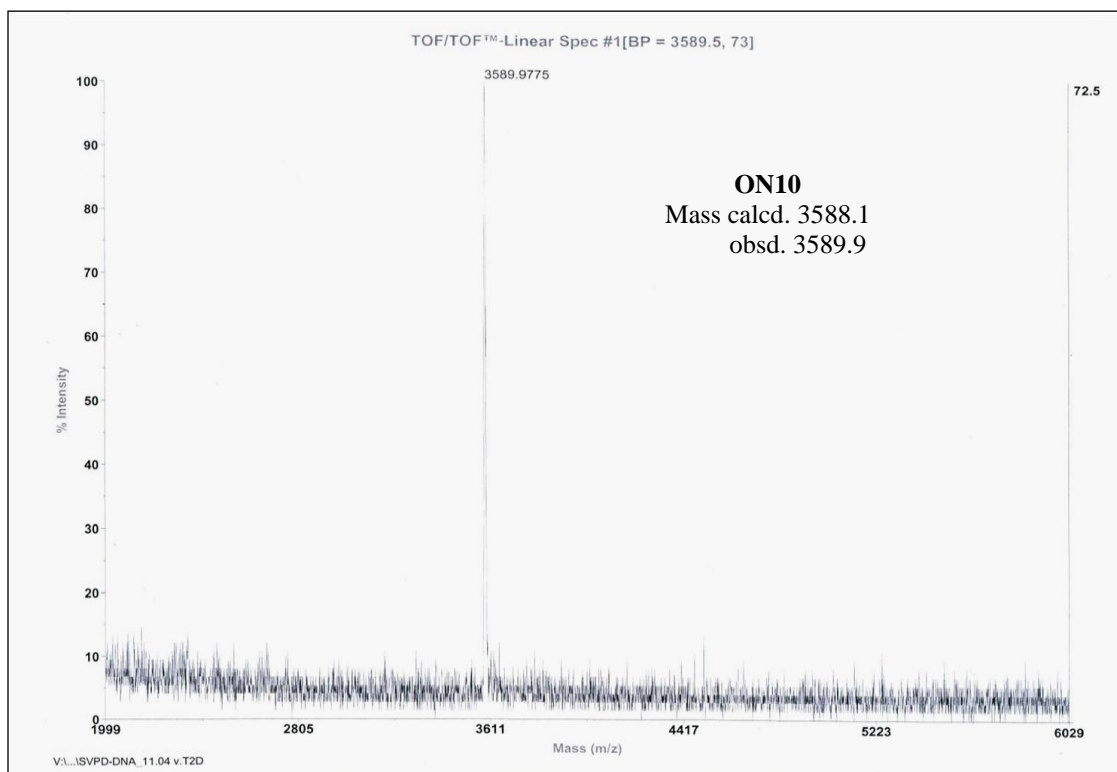
Chapter 2



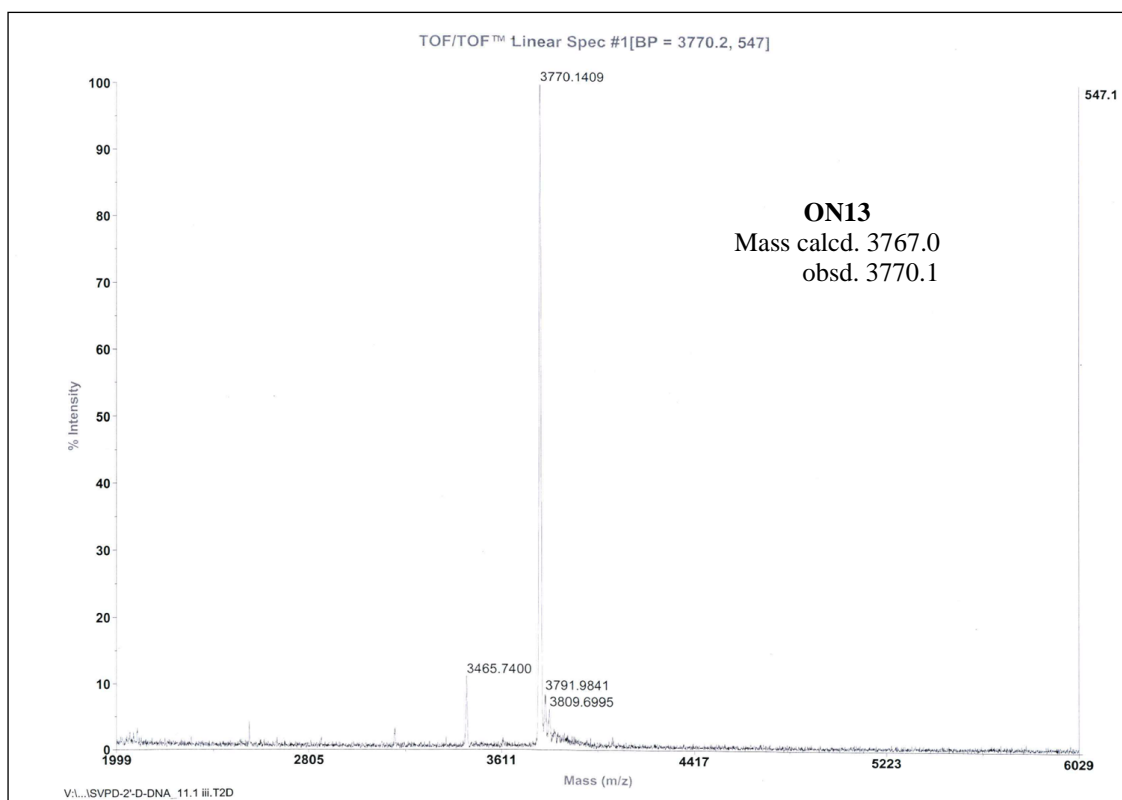
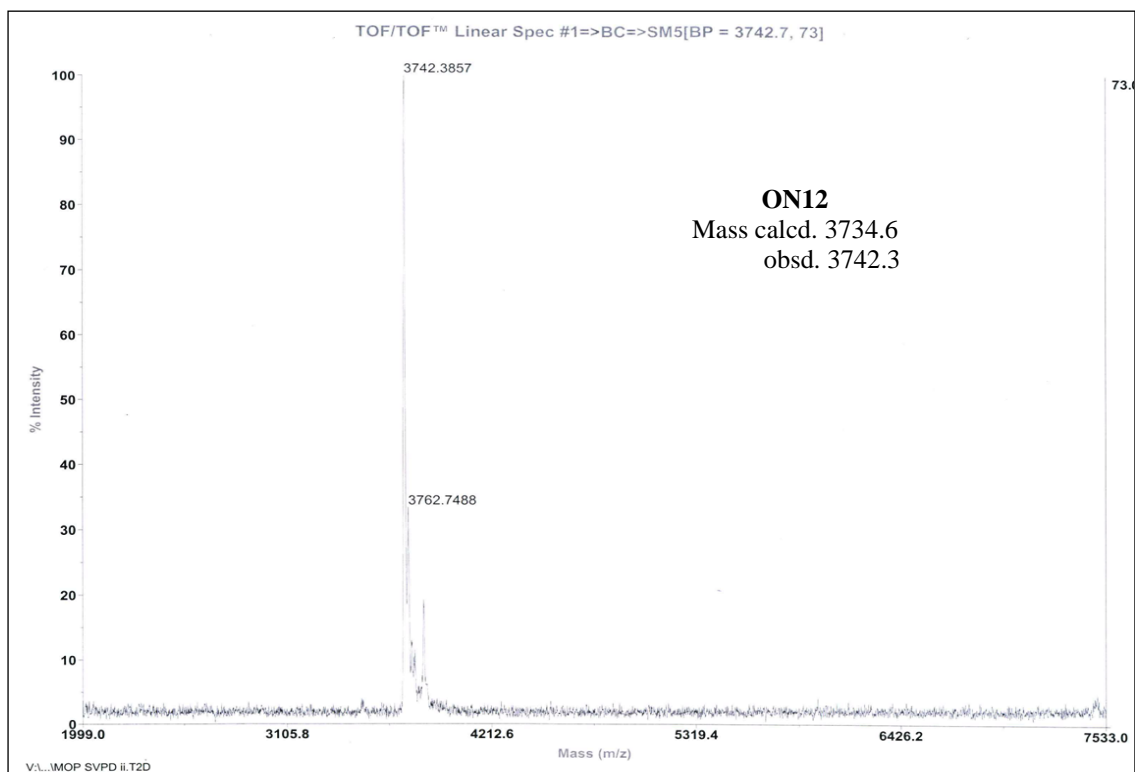
Chapter 2



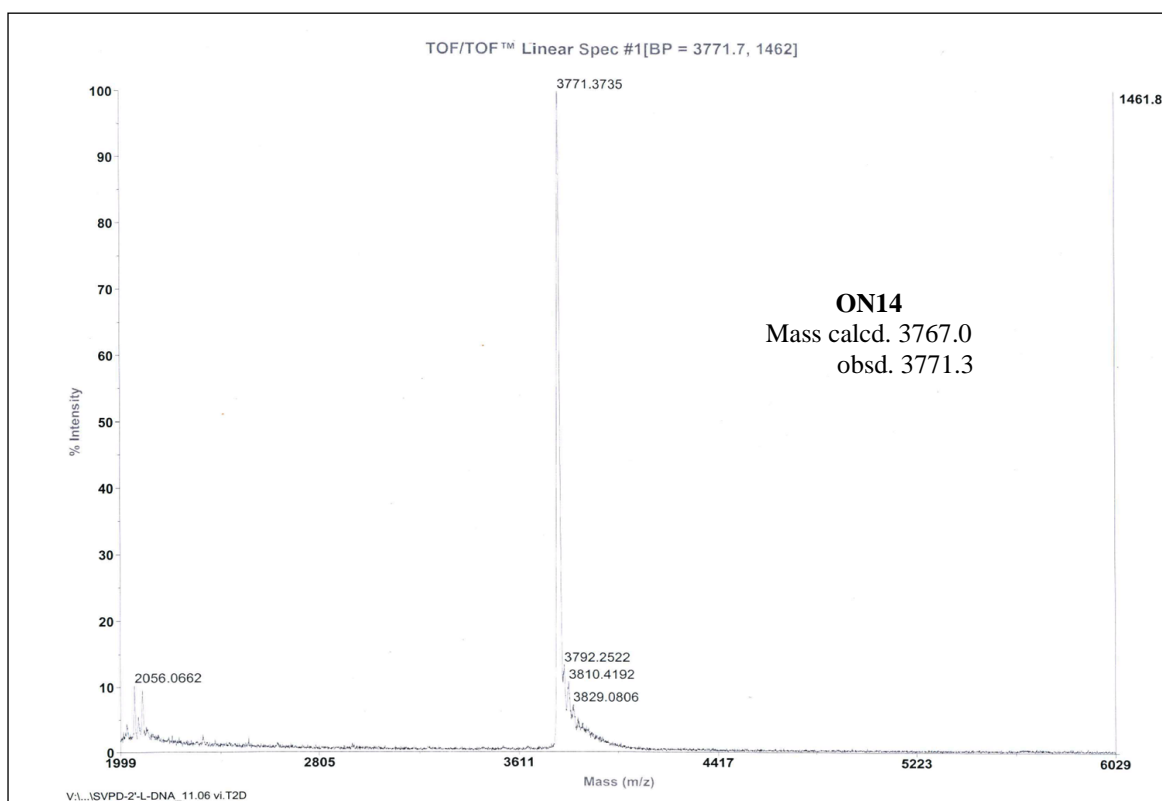
Chapter 2



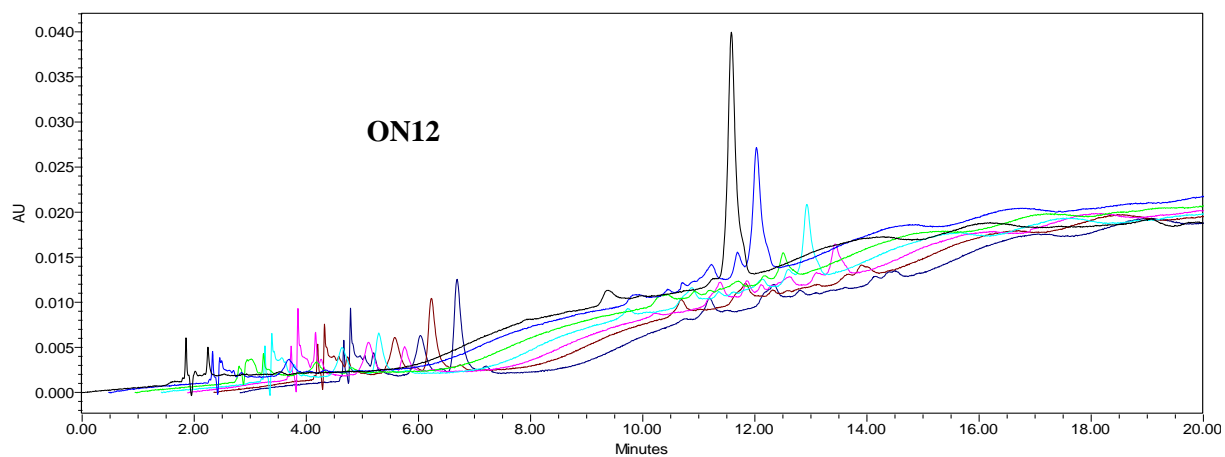
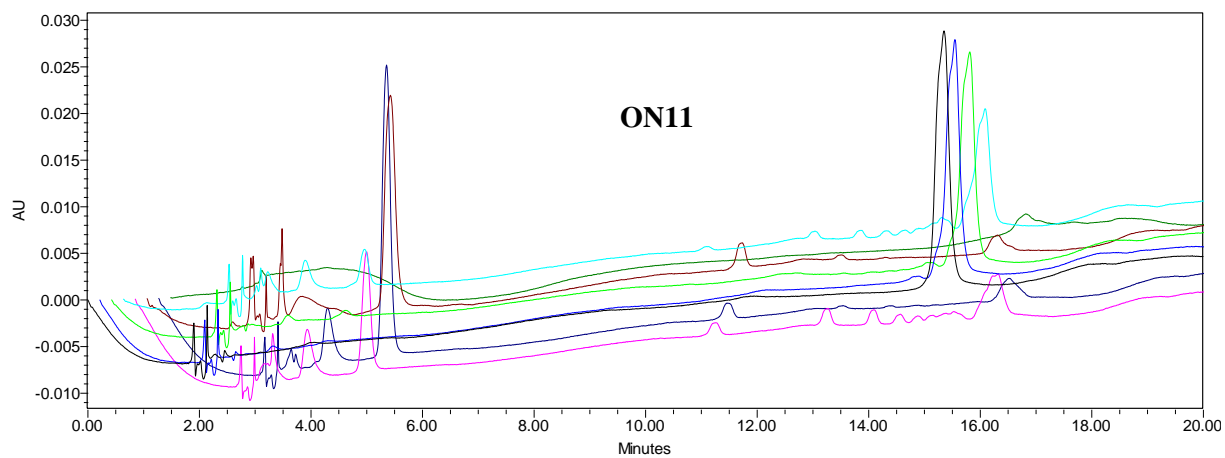
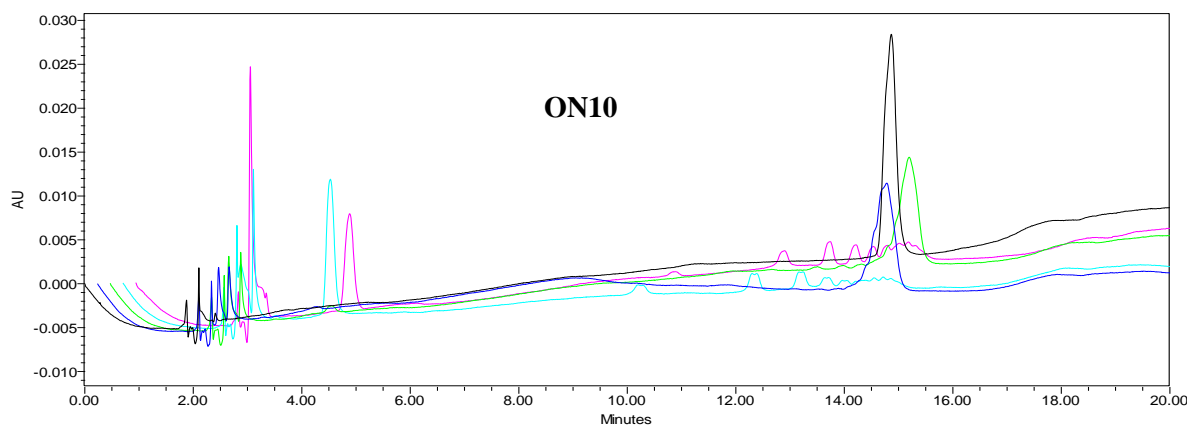
Chapter 2

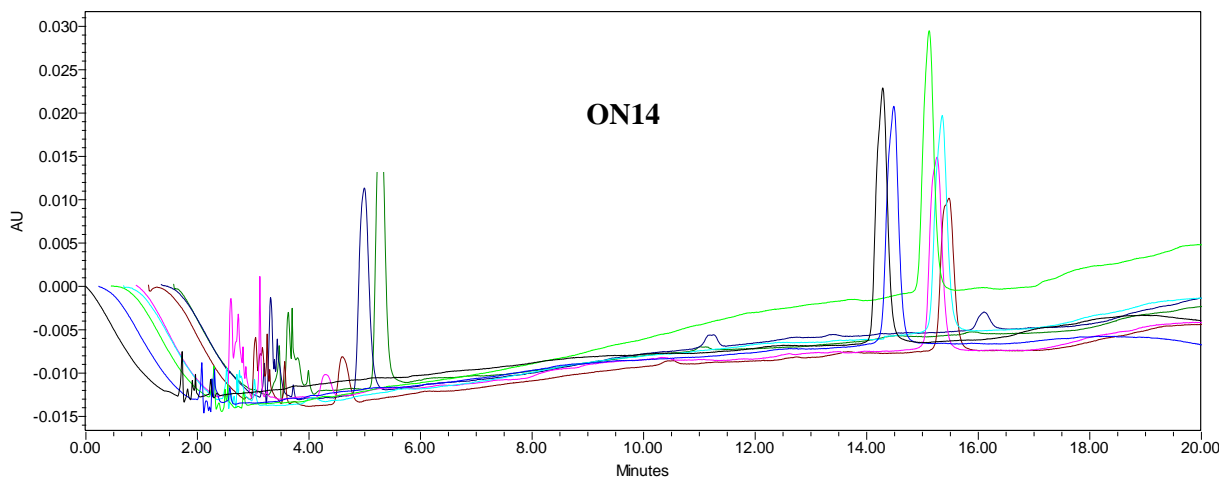
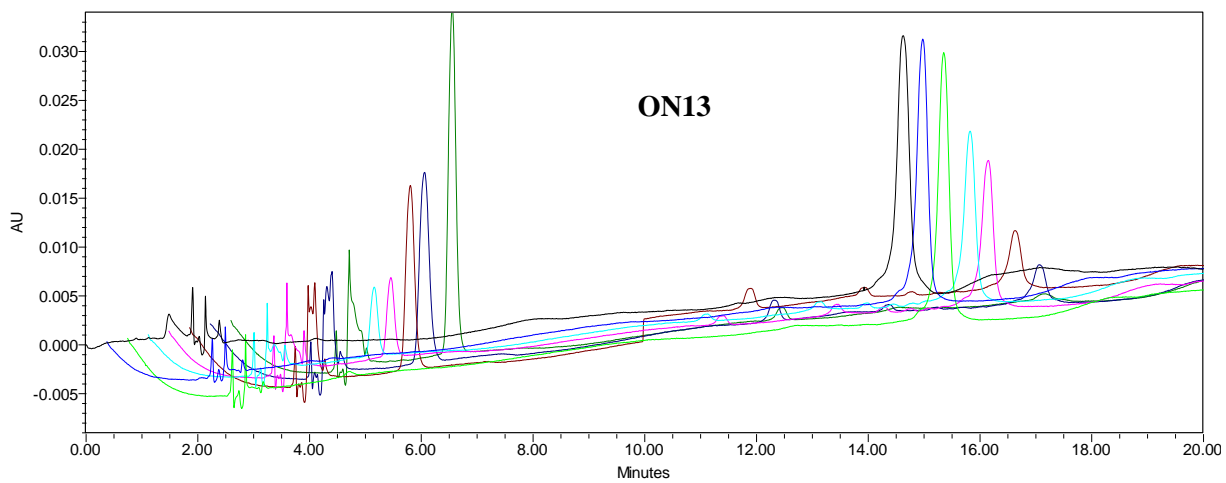


Chapter 2



SVPD digestion- Overlay of HPLC chromatograms





Section 2B: Design and synthesis of 3'/5'-ends serinol capped oligonucleotides and 2'-(ω -O-methylserinol) guarded OMe-RNA/DNA mixmers for efficient splice correction

2B.1 Introduction

An ideal antisense nucleotide (AON) should have high affinity and high specificity to the target, resistance to nuclease attack, low toxicity, and a relatively low costing synthetic route. Chemists designed many chemical modifications which address most of these required elements for an ideal AON, leading to the numerous modified nucleic acids (as discussed in the preceding chapter). The most important factors dictating the potency of the chemical modifications are high binding affinity to the target and high enzymatic stability in biological systems. The chemical modifications that render protection to the AONs against enzymatic hydrolysis can be viewed as mainly of two types. One is a non-nucleosidic/nucleosidic chemical modification which is incorporated into AONs as an additional insertion mostly at the 5' or 3' end or both (top panel, Figure 8a-g), to protect AONs from degradation by exonucleases. The second one is a chemical modification of a nucleoside unit that is incorporated into the AON as a substituent (like 2'-O-alkyl derivatives, bottom panel, Figure 8h, i), to achieve protection against endonucleases.

Zamecnik and Stephenson's discovery of inhibition of Rous sarcoma virus replication by short oligodeoxynucleotide marked the birth of antisense technology.²⁸ In their pioneering work, they found that the efficacy of oligodeoxynucleotides was improved by capping the 3'/5'-ends as carbamates (Figure 8a), by reaction of 3'/5'-OH with phenylisocyanate,²⁹ which reduces the susceptibility of oligodeoxynucleotides to enzymatic degradation. Tennant *et al.* had shown earlier in 1973, that nuclease resistant 2'-OMe-poly(A) exhibited higher inhibitory efficacy of over Poly(A) on murine oncornavirus production in tissue culture.³⁰ These results along with the observation of long hairpin structures that block exonuclease attack inspired researchers to cap the ends of AONs to improve their potency in miRNA targeting.³¹

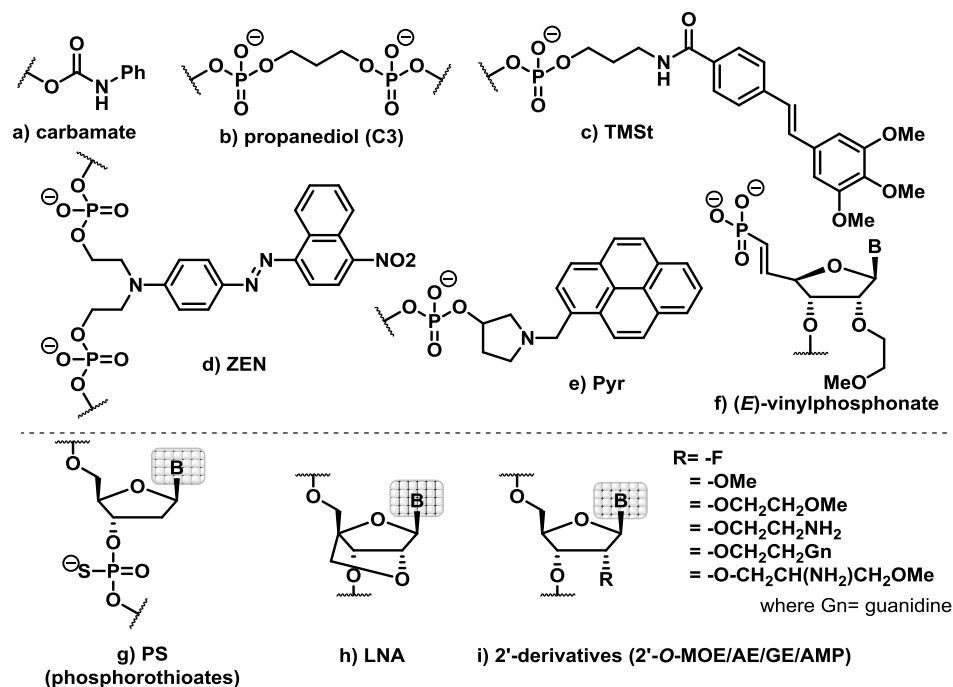


Figure 8 3'/5'-exonuclease or internal endonuclease resistant chemical modifications.

Lennox *et al.* screened non-nucleotide modifiers for their utility in steric blocking antisense applications and reported that N,N-diethyl-4-(4-nitronaphthalen-1-ylazo)-phenylamine (“ZEN”, Figure 8d) was a potent miRNA inhibitor *in vitro*, in combination with the 2'OMe RNA backbone.^{31a} The ZEN modification not only blocked the exonuclease activity but also provided the desired increase in binding affinity to the target miRNA, whereas a simple C3-spacer (Figure 8b) did not show any appreciable inhibitory activity. Recently, Blechinger *et al.* screened several short 2'-OMe RNAs (14-mer) having different capping units at the 5'- and 3'-ends as selected miRNA inhibitors.^{31b} They concluded that 2'-OMe RNAs whose 5'-ends were capped with a trimethoxystilbene residue (TMSt, Figure 8c) or a pyrene derivative (Figure 8e) and 3'-ends with three 2'-fluoro-2'-deoxynucleotides (Figure 8i, R=F) were potent inhibitors of target miRNA.

The chemistry of antisense AONs has progressed immensely over the last 4 decades when several analogs such as phosphorothioates (PS),³² peptide nucleic acids (PNA)³³ and morpholino nucleic acids (PMO)³⁴ were introduced as linkage modifiers and 2'-O-alkyl derivatives^{2a} such as 2'-OMe and 2'-O-methoxyethyl (MOE),^{2a} 2'-F,³⁵ ANA,³⁶ LNA³⁷ etc.

as sugar modifiers, for endowing nuclease resistance to oligomers as well as allowing increased efficiency of complexation with complementary DNA/RNA.^{32, 38}

Though the promising 2'-modifications like 2'-OMe, 2'-F, 2'-O-MOE and LNA are known to increase the duplex stability, the endonuclease protection offered by these modifications against nuclease degradation/hydrolysis is not enough for antisense applications.^{31a} The phosphorothioates (PS), where one of the non-bridging oxygen atom of internucleosidic phosphate backbone is replaced by a sulfur atom, have long been known to protect the AONs from nuclease degradation.³⁹ Unfortunately, the PS linkages destabilize the duplex. In order to increase the potency of AONs, PS linkages are used in combination with other chemical modifications which render increased duplex stability.^{39b, 40} In spite of being diastereomeric mixtures at each phosphorus atom, the PS linkages have not yet found replacement because of their favorable pharmacological properties such as increasing half-life and improved binding to serum proteins *in vivo*, allowing greater availability of AONs to biological targets.⁴¹ Several of the chemistries are being mixed in the recent years to gain maximum advantages in terms of reducing off-target effects, increasing specificity and potency of the AONs in various strategies such as RNase-H dependent antisense,^{32, 38, 42} siRNA,^{3b, 43} miRNA^{31b, 44} or splice switching antisense applications.⁴⁵ Recently Lima *et al.* emphasized the importance of end capping in AONs for siRNA therapeutics (5'-end, with vinylphosphonate) in addition to PS linkages in the backbone.^{3b} During their investigation of potential siRNA therapeutics, they capped the 5'-end of the AON with a metabolically stable (*E*)-vinylphosphonate which is iso-stereoelectronic to the natural phosphate group, whose presence is a must for siRNA action *in vivo*.^{3b} The recent literature again points out the necessity to protect the 3'-/5'- ends or both by enzymatically stable capping of AONs.^{3b, 31}

It has emerged that among the plethora of modified AONs currently under evaluation, the promising AONs have some undesirable drawbacks, e.g., phosphorothioate AONs or OMe/LNA mixmers show non-sequence-specific effects due to non-specific binding to untargeted proteins⁴⁶ or due to mismatched non-target recognition due to very high duplex stability.⁴⁷ The enzyme resistant phosphorothioate AONs are a mixture of

diastereomers at every linker phosphorus atom and the separation of diastereomers is not easy.⁴⁸ Such AONs also show reduced binding efficiency to RNA. The enzyme resistant LNA analogs^{15b, 49} such as c-OMe or c-Et also require several synthetic steps and separation of diastereomers during their synthesis. These shortcomings are indicative of the pressing need for efficient AON analogs that employ relatively simple chemistry and are chirally homogeneous but yet have fewer toxic off-target effects and have higher efficiencies.

2B.2 A simple design and rationale

In the previous section, we have discussed synthesis and properties of our designed 2'-*O*-(*R/S*-2-amino-3-methoxypropyl) (2'-*O*-*R/S*-AMP) modified nucleic acids which were proved to exhibit the synergistic characteristics of 2'-*O*-MOE and 2'-*O*-aminopropyl substitutions in a stereospecific manner.²⁰ Our *R/S*-AMP substitutions retained the duplex stability like 2'-*O*-MOE-containing AONs upon binding to the target RNA, and in addition provided the nuclease resistance as in 2'-*O*-(2-aminoethyl)- tethered nucleic acids and unlike 2'-*O*-MOE-AONs.⁷ The additional nuclease-resistance property endowed by our 2'-*O*-*R/S*-AMP modifications could be explained similar to any other 2'-*O*-cationic tethers, where a protonable functional group displaces divalent metal cations at the active site of the nuclease, thereby restricting its specific function of phosphodiester hydrolysis.⁷

Encouraged by the above results, along with the literature-reported stability against endo-nuclease offered by C3 spacer during the investigation of stable hairpin structures,⁵⁰ we designed a universal serinol cap for both 3' and 5'-ends to impart enzymatic stability to the AONs. Chemically, our design is a simple three carbon spacer which is spatially equivalent to the three carbons from the sugar unit that are involved in the internucleosidic backbone in DNA/RNA (Figure 9a),⁵⁰ but carrying the additional 'amine' functional group which is anticipated to play a crucial role in protection of the AONs from enzymatic degradation (Figure 9b).

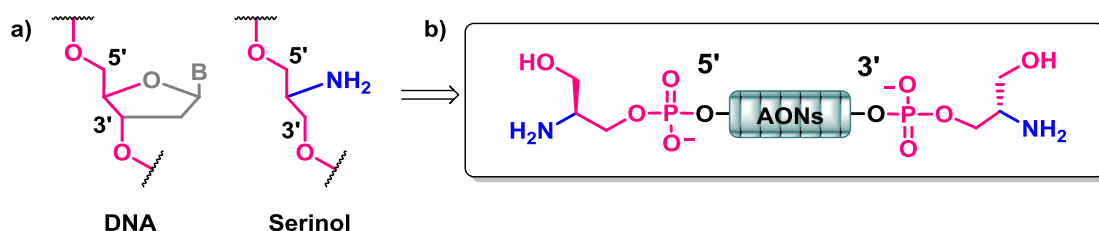


Figure 9 Design of serinol as a universal cap to improve the enzymatic stability of AONs.

To gain the maximum advantage of our 2'-*O*-*R/S*-AMP modifications towards the development of potential AONs we chose the AMP-tether chirality as '*R*', which offered better protection against SVPD than its counterpart *S*-AMP tether, and thymine as a nucleobase in the nucleoside rather than uracil since thymine C5-methyl group was proved to stabilize the duplex.⁵¹ A simple synthetic route to achieve the capping phosphoramidite from L-serine and a similar synthetic route to that presented in the preceding section for 2'-*O*-*R/S*-AMP-uridine derivative was accomplished. To check the stability offered by the serinol capping units, we examined the stability of serinol capped- DNA and 2'-OMe oligonucleotides against SVPD digestion. 2'-OMe RNAs which are known to stabilize the duplex with the target RNA were capped with two serinol units at both 5'- and 3'- ends and the 2'-*O*-*R*-aminomethoxypropyl- ribothymidine units were also substituted at the pre-defined positions. Our designed OMe-mixmers, consisting of both nucleosidic and non-nucleosidic protection derived from L-serine, were shown to have potential antisense applications.

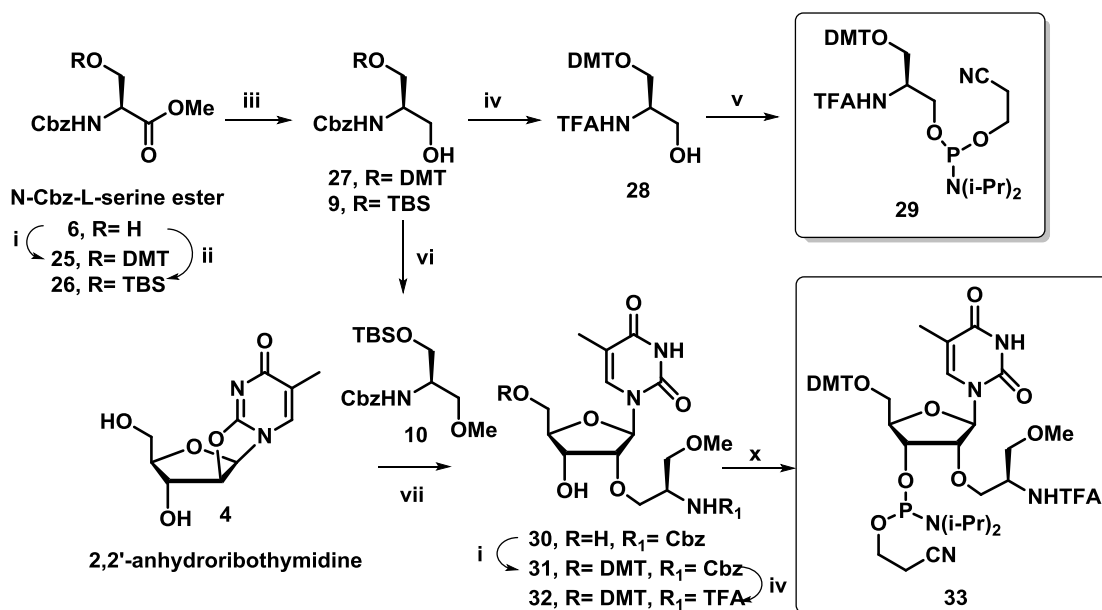
2B.3 Synthesis of amidite derivatives and oligonucleotides

The synthesis of the universal end-capping monomer **29** and the 2'-*O*-*R*-AMP-thymidine monomer (T^{R-AMP}) **33** is outlined in Scheme 6. The primary hydroxyl group of serine derivative **6** was protected as TBS (**25**) or DMT (**26**) ethers, followed by ester reduction to produce unsymmetrized diols **27** and **9**, carrying respective protecting groups. Compound **27** was then hydrogenated to get a free amine, which was again protected as the TFA derivative (**28**) and converted into the universal capping amidite **29**, compatible with solid phase phosphoramidite oligomer synthesis. Alcohol **9** was methylated (**10**) and was

Chapter 2

used to functionalize ribothymidine to get the thymidine derivative **30** via opening of 2,2'-anhydro ring of 2,2'-anhydroribothymidine (**4**), as discussed in the previous Section A. The primary hydroxyl group in **30** was protected as its DMT derivative (**31**), the primary amino group protection was changed to TFA (**32**) and **32** was finally converted to T^{R-AMP} amidite **33** compatible with solid phase phosphoramidite oligomer synthesis.

Scheme 6 Synthesis of universal capping amidite **29** and T^{R-AMP} amidite **33**



Reagents and conditions: (i) DMT-Cl, cat.DMAP, dry pyridine (ii) TBS-Cl, imidazole, dry DCM (iii) NaBH₄, MeOH, 0 °C, rt, 78% over two steps for **27** and 87% over two steps for **9** (iv) (a) H₂, Pd/C, 65 psi, MeOH, rt, 4 h (b) CF₃COOEt, NEt₃, MeOH, rt, 86% over two steps (v) 2-cyanoethyl N,N-diisopropylchlorophosphoramidite, DIPEA, dry DCM, 3 h, 85% (vi) Ag₂O, MeI, dry ACN, rt, 12 h, 91% (vii) **10**, BF₃.Et₂O, dry DMA, 130 °C, 8 h, 65% (viii) 2-cyanoethyl N,N-diisopropylchlorophosphine, DIPEA, dry DCM, 3 h, 82%.

The phosphoramidites **29** and **33** were incorporated into sequences using increased coupling time (6 min) to yield the modified sequences, listed in Table 3. All the sequences were purified by reverse phase HPLC and were characterized by MALDI-ToF mass analysis (Table 4). The sequence **ON1** is a control sequence without any end capping. Sequences **ON15** and **ON16** correspond to the single and double end capping (by employing monomer **29**) at both 3'- and 5'-ends. **ON17** is the 2'-OMe derivatised RNA

Chapter 2

sequence. **ON18** and **ON19** are the single and double serinol (**29**) capped OMe-RNA sequences, respectively. In the sequence **ON20**, four 2'-OMe-uridine residues are replaced by T^{R-AMP} derivatives (**33**) at defined positions. The single and double capping of the sequences with the serinol derivative **29** was performed to assess the resistance offered by the capping to hydrolysis by phosphodiesterase enzyme.

2B.4 UV- T_m studies and SVPD digestion of the capped oligonucleotides

Table 4 The sequences synthesized, their MALDI-ToF mass analyses and UV- T_m values with complementary RNA and RNA with single mismatch (top panel); commercially purchased sequences used in the study (bottom panel)

Sequence/Code ^a	Mass Calcd/Obsd	UV- T_m °C ^b cRNA/mmRNA
CCTCTTACCTCAGTTACA ON1	5366.9/5367.2	56.6/50.4
SCCTCTTACCTCAGTTACAS ON15	5672.9/5690.3	58.3/50.4
SSCCTCTTACCTCAGTTACASS ON16	5979.0/6002.0	58.4/48.7
ccucuuaccucaguuaca ON17	5823.0/5824.4	77.0/66.2
Sccucuuaccucaguuaca S ON18	6129.0/6151.4	78.2/67.9
SSccucuuaccucaguuaca SS ON19	6435.1/6453.0	76.8/65.4
SSccxcu x accxcagu x aca SS ON20	6786.6/6787.8	76.4/64.0

c ^L uc ^L uA ^L c ^L uc ^L A ^L gu ^L aC ^L a ON21		
c ^L T ^L cuC ^L uC ^L aC ^L caT ^L ugA ^L cA ^L a ON22		
ugu aac uga ggu aag agg (cRNA)		
ugu aac ugc ggu aag agg (mmRNA)		

^aa/c/g/u are RNA or 2'-OMeRNA nucleotides; **S** is a serinol capping unit; x= T^{R-AMP} is a 2'-O-(2-amino-3-methoxypropyl)ribothymidine unit; C^L , A^L and T^L are the conventional locked nucleotides
^bUV- T_m values were measured by using 1 μ M sequences with 1 μ M cDNA/cRNA in sodium phosphate buffer (0.01M, pH 7.2) containing 100 mM NaCl and are averages of three independent experiments. (Accuracy is ± 0.5 °C).

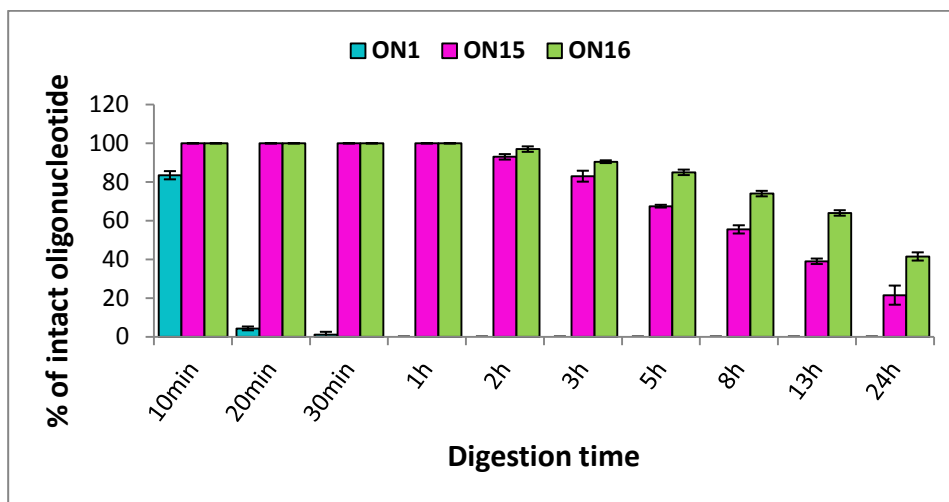


Figure 10 SVPD digestion of 5'/3'-end capped DNA oligonucleotides (ON15 and ON16) compared to native DNA (ON1). Experiments were done in duplicate and the error bars (standard deviations) are indicated. (Please see the experimental section for the conditions used).

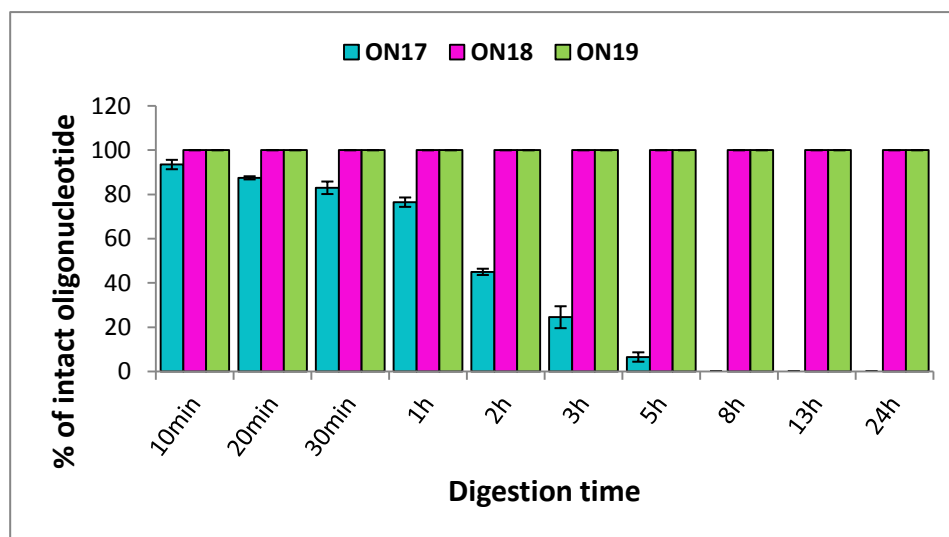


Figure 11 SVPD digestion of 5'/3'-end capped 2'-OMe RNA oligonucleotides (ON18 and ON19) compared to native 2'-OMe RNA (ON17). Experiments were done in duplicate and the error bars (standard deviations) are indicated. (Please see the experimental section for the conditions used).

We studied the protection of these synthesized capped oligomers against hydrolytic cleavage *in vitro* and compared the results with unmodified ON1 and ON17. We digested these sequences with SVPD under conditions reported earlier.^{15e, 20} The products of the digestion were analyzed by RP-HPLC and percent intact AON was plotted against time to

understand the degradation pattern for all the oligomers (Figures 10 and 11). After digestion with SVPD for 24h, both the capped DNA sequences **ON15** and **ON16** were found to resist hydrolysis to some extent and were found to be degraded by 70% and 50%, respectively, while the native **ON1** was completely digested after ~30 min. (Figure 10). The protective effect was more pronounced in 2'-OMe RNAs, where the terminal serinol units were found to be able to completely guard the 2'-OMe-RNA oligomers (**ON18** and **ON19**) against hydrolytic cleavage by SVPD for up to 24 h compared with the uncapped OMe-RNA sequence **ON17** (Figure 11).

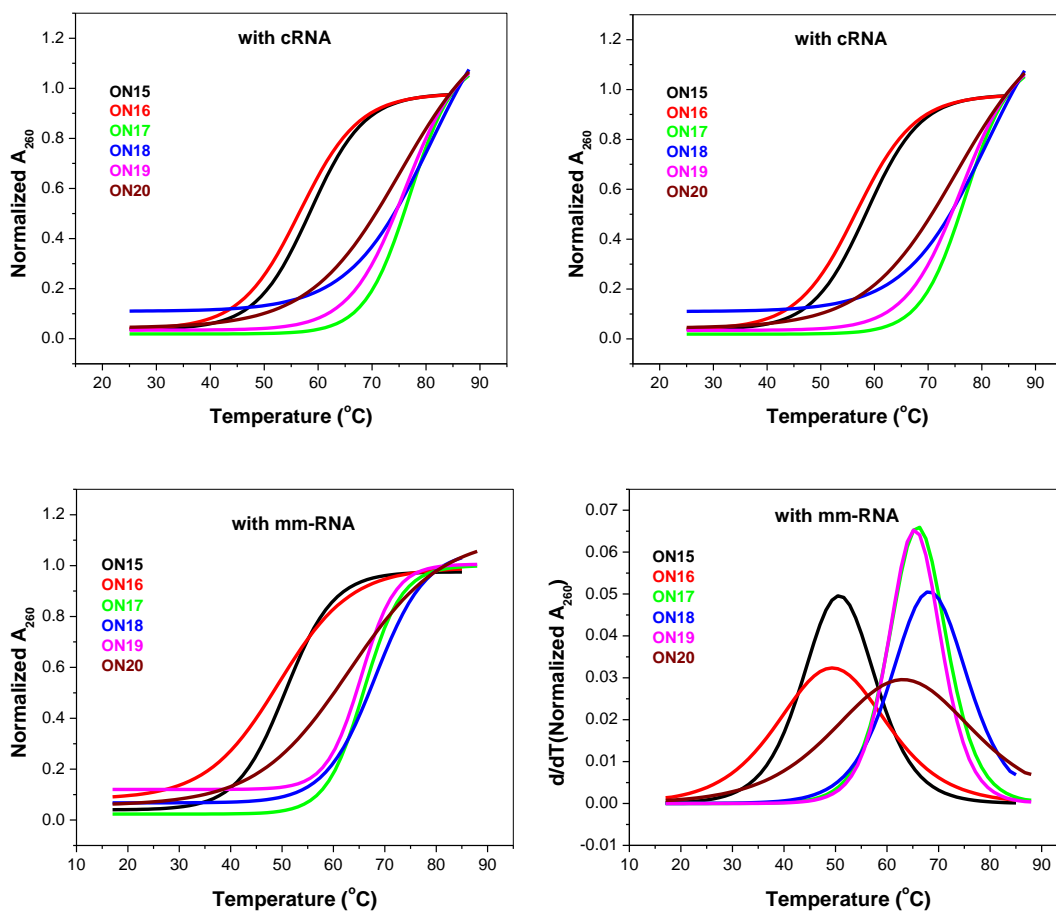


Figure 12 Sigmoidal transitions of thermal denaturation experiments and their first order derivative curves of modified oligonucleotides with cRNA and mmRNA. (UV- T_m values were measured by annealing 1 μ M sequences with 1 μ M cRNA/mmRNA in 10 mM sodium phosphate buffer, pH 7.2 containing 100 mM NaCl, averages of three independent experiments, and are accurate to ± 0.5 $^{\circ}$ C).

Having confirmed the enzymatic stability conferred to these sequences by end capping with serinol and 2'-OMe introduction, we studied the effect of these modifications on binding affinity by UV- T_m studies with complementary cRNA and single base mismatched mmRNA (Table 4). It was observed that the end capping as well as the incorporation of T^{R-AMP} units have negligible effect on the binding efficiency of the 2'-OMe oligomer. The mismatch discrimination was found to be slightly better or equal to that of the 2'-OMe oligomer (Table 4), indicating that the amino pendant groups do not lead to any additional non-specific electrostatic interactions with target RNA. The stabilization effect of T^{R-AMP} introduction is comparable to that of OMe substitution.

2B.5 Splice correction activity of 3',5'-serinol- and 2'-(ω -O-methylserinol) guarded 2'-OMe-mixmers

For biological experiments, to harness the maximum benefits of our designed modifications, we chose the sequences **ON19** and **ON20**, each having two serinol caps at both 3'- and 5'-ends, which were found to be competent at duplex formation as well as resisting enzymatic hydrolysis. Cell lines expressing luciferase pre-mRNA, interrupted by an aberrantly spliced β -globin intron, HeLa pLuc705, were used to monitor the splice-switching activity of the modified oligonucleotides. Various AONs have been evaluated using these cell lines by measuring induced luciferase activity as a tool to verify splice-switching efficiency and specificity.⁵² **ON19** and **ON20** were used in this study in comparison with OMe-RNA (**ON17**), and also with an LNA-OMe mixmer (**ON21**), recently used in steric blocking applications.^{44, 53} The expression of functional luciferase in the form of RLU/ μ g protein expression was studied as an indication of splice-switching efficiency.

The results shown in Figure 13 clearly indicate that the dose-dependent luciferase expression in the cells treated with serinol-end-capped **ON19** was better than that for the uncapped OMe-RNA (**ON17**). The expression was found to be even better for the sequence **ON20**, which comprised four T^{R-AMP} units (i.e. ~20% substitution of U^{OMe} with 2'- O - T^{R-AMP}), in addition serinol end-caps. The expression of corrected fluorescent protein was found to be as good as or even slightly better than with **ON21**, an LNA/OMe mixmer

containing about ~40% LNA substitution. LNA substitution is known to impart very high stability to the duplexes with cRNA (~5 °C/LNA unit).^{37, 54}

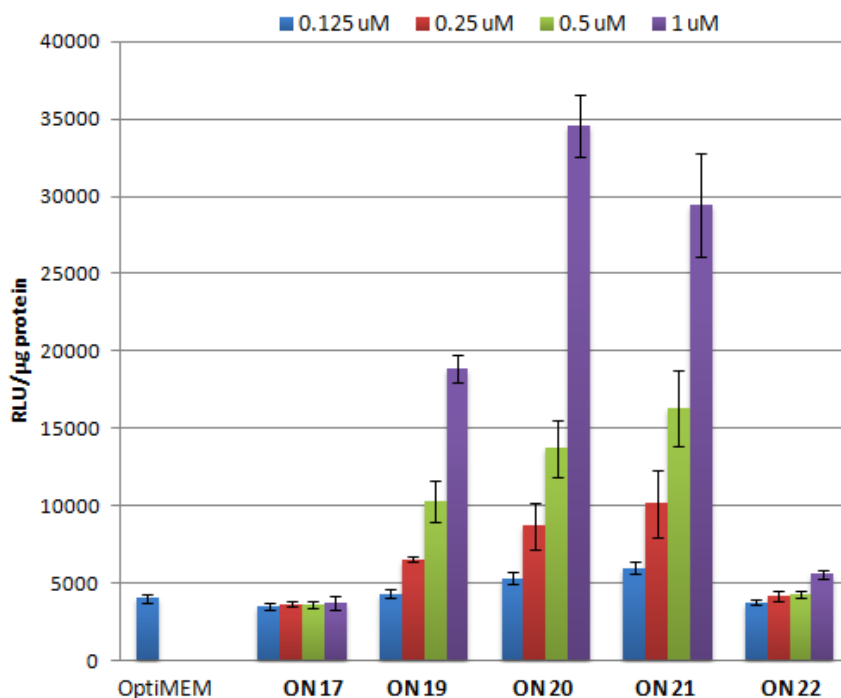


Figure 13 Concentration dependent splice correction activity of oligonucleotides.

In addition, LNA/OMe mixmers have been previously reported to show high nuclear steric blocking activity following transfection whereas OMe-RNA was inactive.⁵³ The present modification, in the sequence **ON20**, does not have the advantage/disadvantage of LNA in the context of additive duplex stability per modified unit but is still able to exhibit the ability to enter the cell nucleus and give high splice correction activity. As is evident from Figure 13, just the end capping as in **ON19** was not enough to obtain good activity but the interspersed four T^{R-AMP} units in **ON20** enhanced the luciferase expression significantly and to a level obtained by as much as seven LNA units. This result is quiet remarkable. This could be due to the balanced duplex stability and protection against enzymatic susceptibility endowed by internal T^{R-AMP} units in addition to the enzymatic stability of end capping. We have also shown by UV-T_m studies (Figure 12), that the single

mismatch discrimination is as high as that shown by OMe-RNA oligomer while forming a duplex with the target RNA and therefore off-target effects may also be expected to be less compared with the LNA-OMe mixmer sequences, which sometimes tolerate mismatched bases in the target RNA.⁴⁷ The oligomer **ON20** with 8 additional amino groups, which partially neutralizes the negative charge of the phosphate backbone, was used also in luciferase activity experiments without transfection agent, but no activity was observed (Please see experimental section). This is perhaps not surprising since there is still a net negative charge on the oligomer.

Currently used therapeutic oligomers employ phosphorothioate substitution to circumvent the susceptibility of the oligomers to hydrolytic enzymes and also to enhance cellular and *in vivo* activity. Therefore, it would be interesting to study if our novel strategy of 2'- substitution presented here can be used to improve cell uptake further and possibly even to avoid (or at least reduce) the requirement for phosphorothioate substitution. The results presented here may have implications in other diseases including cystic fibrosis, muscular dystrophies, cancers, and several neurological disorders where there are associated mutations affecting the splicing process.³²

2B.6 Conclusion

- A simple design for an efficient protection of oligonucleotides and the synthesis of the corresponding amidites were achieved in possibly the simplest way.
- To harness the maximum benefit of our serinol-derived modifications, we incorporated our nucleosidic and non-nucleosidic serinol derivatives in 2'-OMe RNAs without disturbing the efficacy or fidelity of duplex formation.
- The 5'- and 3'-end capping of oligonucleotides with serinol provided excellent protection against SVPD digestion. The single and double cappings with serinol at both 5'- and 3'-ends offered 30% and 50% protection to the DNA oligonucleotides respectively. Further, the single and double cappings with serinol at the ends provided 100% protection to the 2'-OMe oligonucleotides. The difference in the extent of

protection could easily be attributed to the inherent protection of 2'-OMe RNAs against nucleases.

- We showed that the 3'- and 5'- capped 2'-OMe-AON with ~20% evenly dispersed modified T^{R-AMP} units is as effective as an LNA-OMe mixmer containing ~40% LNA for antisense applications in steric blocking splice correction of an aberrant β -globin gene, i.e., a much lower degree of modification (in our case, 2'-*O-R-AMP*) was shown to be as effective as LNA-OMe mixmer containing a much higher degree of modification (LNA).

2B.7 Experimental Section

General

All the non-aqueous reactions were performed under the inert atmosphere of Nitrogen/Argon and the chemicals used were of laboratory or analytical grade. All solvents used were dried and distilled according to standard protocols. TLCs were performed on pre-coated silica gel GF254 sheets (Merck 5554). All reactions were monitored by TLC and usual work-up implies sequential washing of the organic extract with water and brine followed by drying over anhydrous sodium sulfate and evaporation of the organic layer under vacuum. Column chromatographic separations were performed using silica gel 100-200 mesh (Merck) or 230-400 mesh (Merck) and using the solvent systems EtOAc/Petroleum ether or MeOH/DCM. ^1H and ^{13}C NMR spectra were obtained using Bruker AC-200, AC-400 or AC-500 NMR spectrometers. The chemical shifts are reported in delta (δ) values and referred to internal standard TMS for ^1H . HRMS-mass spectra were recorded on a Finnigan-Matt mass spectrometer. cDNA sequences were synthesized on Bioautomation Mer-Made 4 DNA synthesizer using standard β -cyanoethyl phosphoramidite chemistry. DNA sequences were also analyzed and purified under the same conditions with an increasing gradient of acetonitrile in 0.1N triethylammonium acetate, pH 7.0. The MALDI-TOF spectra were recorded on AB Sciex TOF/TOFTM series explorerTM 72085 instrument; THAP (2, 4, 6-trihydroxyacetophenone) was used as the matrix. Thermal denaturation experiments were performed on Varian Cary-300 UV-Vis spectrophotometer fitted with a peltier-controlled temperature programmer and a water circulator, at the temperature ramping rate 0.5 °C/min and the absorbance was recorded at 260 nm for every 0.5 °C rise in temperature. OMe-RNA/LNA and OMe-RNA/LNA/scr oligomers were obtained from RiboTask ApS, Odense, Denmark.

SVPD digestion

Stability of the capped oligonucleotides towards the exo-nuclease SVPD (snake venom phosphodiesterase) was analyzed by RP-HPLC. 7.5 μM of oligonucleotide in 300 μL of Tris-HCl buffer (pH=7.5, 10 mM Tris-HCl, 8 mM MgCl_2) was incubated at 37 °C for 15

min. SVPD 5 ng/ μ L was added to the oligonucleotide incubated at 37 °C and aliquots of 20 μ L were removed at time intervals of 10 min, 20 min, 30 min, 1 h, 2 h, 3 h, 5 h, 8 h, 13 h and 24 h. Aliquots were kept at 90 °C for 2 min and analysed by RP-HPLC using an increasing gradient of acetonitrile in triethylammonium acetate (A: 5% acetonitrile and B: 30% acetonitrile in 0.1 N triethylammonium acetate, pH 7.0). The percent intact oligonucleotide (based on the peak area) was plotted against the time.

Splice correction assay

The splicing redirection assay was carried out by the method previously reported⁵⁵ but using the transfection agent Lipofectamine 2000 as delivery agent. Each experimental measurement was carried out in triplicate.

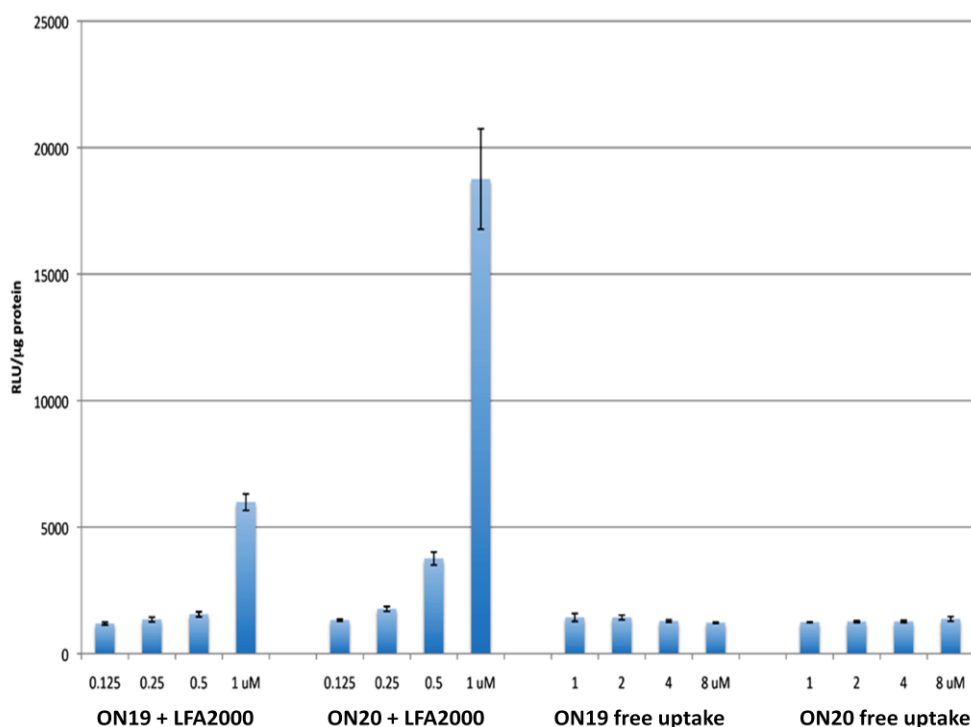
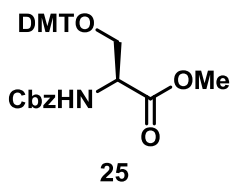


Figure 14 Concentration dependent relative luciferase expression of OMe-RNASb and OMe-RNA4-dTSb oligonucleotides with and without transfection agent Lipofectamine 2000.

Methyl *N*-[(benzyloxy)carbonyl]-*O*-dimethoxytrityl-*L*-serinate, 25:

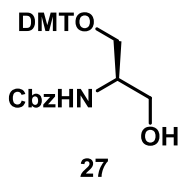
N-Cbz-*L*-serine methyl ester **6** (19.7 mmol, 5.0 g) was dissolved in dry pyridine (25 mL) and DMT-Cl (21.6 mmol, 7.3 g) and catalytic amount of DMAP (~30 mg) were added. Reaction mixture was kept for stirring at room temperature for 4 h. Pyridine was removed under reduced pressure and the residue was diluted with EtOAc. 10% aqueous NaHCO₃, water and



brine solution wash were given to the organic layer. The organic layer was dried over anhydrous Na₂SO₄ and concentrated to dryness. The crude compound was column purified (eluted in 20% EtOAc in petroleum ether) to result in **25** as a pale thick liquid in 91% yield. ¹H NMR (CDCl₃, 200 MHz): δ 3.34–3.39 (dd, 1H, J = 2.65 and 9.09 Hz), 3.54–3.60 (dd, 1H, J = 2.78 and 9.09 Hz), 3.76 (s, 9H), 5.11 (s, 2H), 5.67 (d, 1H, J = 8.59 Hz), 6.78–6.82 (m, 4H), 7.22 (m, 5H), 7.34 (m, 9H); ¹³C NMR (CDCl₃, 50 MHz): δ 52.0, 54.2, 54.8, 63.3, 66.6, 77.2, 85.7, 112.9, 126.6, 127.6, 127.68, 128.2, 128.9, 129.6, 135.1, 35.2, 136.0, 144.1, 155.6, 158.3, 170.8; **HRMS (EI)**: Mass calculated for C₃₃H₃₃O₇NNa (M+Na) 578.2149, found 578.2148.

Benzyl (*R*)-(1-[(dimethoxytrityl)oxy]-3-hydroxypropan-2-yl)carbamate, 27:

Compound **25** (18.0 mmol, 10 g) was dissolved in MeOH (500 mL) and NaBH₄ (72 mmol, 2.7 g) was added fraction-wise at 0 °C for a period of 40 min and the mixture was kept for stirring at room temperature for another 6–8 h. Excess NaBH₄ was quenched with saturated NH₄Cl solution, followed by the removal of MeOH under reduced pressure. The crude reaction mixture was



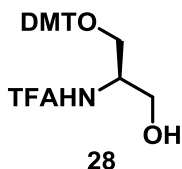
extracted with EtOAc. The organic extract was washed with brine solution and dried over anhydrous Na₂SO₄. EtOAc was removed under reduced pressure to give the crude product and was purified through column chromatography (eluted in 25–30% EtOAc in petroleum ether) to yield **27** as a colorless liquid in 85% (7.8 g). ¹H NMR (CDCl₃, 200 MHz): δ 2.39 (bs, 1H), 3.29 (d, 2H, J = 3.67 Hz), 3.73–3.79 (m, 8H), 5.10 (s, 2H), 5.32 (d, 1H, J = 7.58 Hz), 6.79–6.84 (m, 4H), 7.15–7.40 (m, 14H); ¹³C NMR (CDCl₃, 50 MHz): δ 52.5, 54.9,

Chapter 2

62.9, 66.6, 86.2, 113.0, 126.7, 127.7, 127.8, 128.3, 129.7, 130.5, 136.2, 144.4, 156.3, 158.3; **HRMS (EI)**: Mass calculated for $C_{32}H_{33}O_6NNa$ (M+Na) 550.2200, found 550.2198.

(R)-N-(1-[(dimethoxytrityl)oxy]-3-hydroxypropan-2-yl)-2,2,2-trifluoroacetamide, 28:

Compound **27** (8.7 mmol, 4.6 g) was dissolved in MeOH (20 mL) followed by the addition

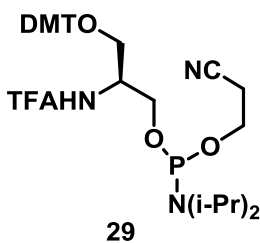


of 10% Pd-C (10% w/w, 0.46 g). The reaction mixture was subjected to catalytic hydrogenation at 60 psi of hydrogen pressure for 5 h. After TLC analysis, the reaction mixture was filtered over celite and the removal of methanol *in vacuo* gave free amine. Without further purification, the amine

was subjected to trifluoroacetyl protection. To the crude amine (8.4 mmol, 3.3 g) dissolved in MeOH (30 mL), NEt_3 (12.7 mmol, 1.7 mL) was added. Ethyltrifluoroacetate (10.1 mmol, 1.2 mL) was added to the reaction mixture and the mixture was kept stirring at room temperature for 8h. MeOH was removed on rota evaporator and the reaction mixture was diluted with EtOAc. The organic layer was washed with water and 5% aq. $NaHCO_3$ and the organic layer was dried over anhydrous Na_2SO_4 , concentrated *in vacuo*. The crude compound was subjected to column purification to furnish **28** (eluted in 17% EtOAc in petroleum ether) in 86% yield over two steps. 1H NMR ($CDCl_3$, 200 MHz): δ 2.39 (bs, 1H), 3.37(d, 2H, $J = 4.42$ Hz), 3.68–3.75 (m, 1H), 3.78 (s, 6H), 3.83–3.90 (m, 1H), 4.05–4.14 (p, 1H), 6.82–6.86 (m, 4H), 7.26–7.40 (m, 9H); ^{13}C NMR ($CDCl_3$, 50 MHz): δ 51.4, 55.0, 61.9, 62.1, 77.2, 86.5, 113.2, 126.9, 127.7, 127.9, 129.7, 135.2, 144.1, 158.5; **HRMS (EI)**: Mass calculated for $C_{26}H_{26}O_5NF_3Na$ (M+Na) 512.1655, found 512.1653.

Serinol derived phosphoramidite, 29:

To the compound **28** (2.0 mmol, 1 g) dissolved in dry DCM (12 mL), DIPEA (3.6 mmol,



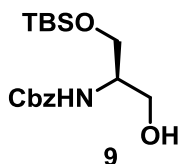
0.6 mL) was added. 2-cyanoethyl-*N,N*-diisopropyl-chlorophosphine (2.4 mmol, 0.5 mL) was added to the reaction mixture at 0 °C and continued stirring at room temperature for 1.5 h. The contents were diluted with DCM and washed with 5% $NaHCO_3$ solution. The organic phase was dried over anhydrous Na_2SO_4 and concentrated.

The residue was re-dissolved in DCM and the compound was precipitated with n-hexane to

obtain corresponding phosphoramidite **33** in 85% yield. ^{31}P NMR (Acetonitrile, D_2O as external standard, 160 MHz): δ 149.17, 149.20; HRMS (EI): Mass calculated for $\text{C}_{35}\text{H}_{43}\text{O}_6\text{N}_3\text{F}_3\text{PNa}$ (M+Na) 712.2734, found 712.2728.

Benzyl (R)-(1-[(tert-butyldimethylsilyl)oxy]-3-hydroxypropan-2-yl)carbamate, 9:

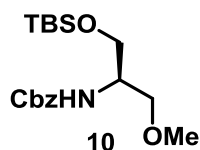
N-Cbz protected-L-serine-methyl ester **6** (39.5 mmol, 10 g) was dissolved in dry DCM (200 mL), followed by the addition of imidazole (98.8 mmol, 6.7 g) and TBS-Cl (47.4 mmol, 7.1 g). The reaction mixture was diluted with DCM and water wash, brine wash were given. The organic layer was dried over anhydrous Na_2SO_4 and solvents were removed under reduced pressure to give the



crude TBS protected ester, which was subjected to NaBH_4 reduction. The residue was dissolved in methanol (500 mL) and cooled to 0 °C, NaBH_4 (150 mmol, 5.6 g) was added fraction-wise over a period of 1 h and then stirring continued at room temperature for another 6 h. Excess NaBH_4 was quenched with saturated NH_4Cl solution, followed by the removal of MeOH under reduced pressure and the product extracted in EtOAc. The organic extract was washed with brine and dried over anhydrous Na_2SO_4 . EtOAc was removed under reduced pressure to give the crude product that was purified by column chromatography (eluted in 15% EtOAc in petroleum ether) to yield **9(R)** in 87% yield (11.6 g) over two steps. ^1H NMR (CDCl_3 , 200 MHz): δ 0.05 (s, 6H), 0.88 (s, 9H), 2.64 (bs, 1H), 3.66–3.84 (m, 5H), 5.11 (s, 2H), 5.38–5.41 (m, 1H), 7.35–7.38 (m, 5H); ^{13}C NMR (CDCl_3 , 50 MHz): δ -5.6, 18.1, 25.7, 53.0, 63.0, 63.4, 66.7, 128.0, 128.4, 136.2, 158.2; HRMS (EI): Mass calculated for $\text{C}_{17}\text{H}_{30}\text{O}_4\text{NSi}$ (M+H) 340.1939, found 340.1945.

Benzyl (R)-(1-[(tert-butyldimethylsilyl)oxy]-3-methoxypropan-2-yl)carbamate, 10:

To a stirred solution of **9** (29.4 mmol, 10 g) and MeI (147.4 mmol, 9.5 mL) was added Ag_2O (73.5 mmol, 16.9 g). The reaction mixture was stirred at room temperature for 12 h, filtered and the filtrate was concentrated under reduced pressure. The crude compound was purified by column chromatography (eluted in 5% EtOAc in petroleum ether) to result **10** in

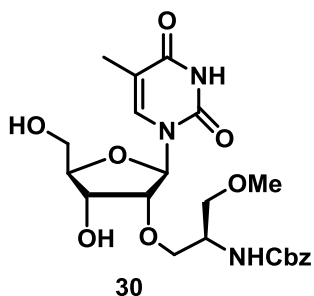


91% (9.4 g) yield. ^1H NMR (CDCl_3 , 200 MHz): δ 0.05 (s, 6H), 0.88 (s, 9H), 3.33 (s, 3H),

3.37–3.77(m, 4H), 3.78–3.86(m, 1H), 5.10 (s, 2H), 7.35–7.38 (m, 5H); ^{13}C NMR (CDCl_3 , 100 MHz): δ -5.6, 18.1, 25.7, 51.5, 58.8, 61.3, 66.6, 70.4, 128.0, 128.4, 136.4, 155.9; **HRMS (EI)**: Mass calculated for $\text{C}_{18}\text{H}_{32}\text{O}_4\text{NSi}$ (M+H) 354.2095, found 354.2089.

2'-functionalization of 2,2'-anhydroribothymidine, **30**:

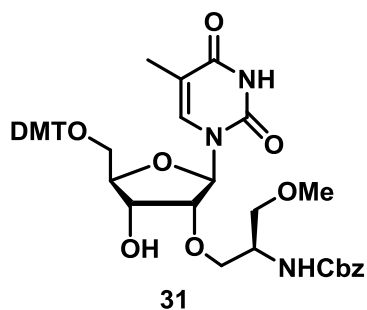
Desiccated 2,2'-anhydrothymidine (6.2 mmol, 1.5 g) was dissolved in dry DMA (10 mL) followed by the addition of $\text{BF}_3\cdot\text{OEt}_2$ (12.5 mmol, 1.5 mL) under argon atmosphere. After 2 min, the $-TBS$ protected silyl ethers **10** (18.7 mmol, 6.6 g) were added and stirring continued at 130 °C for 8 h. DMA was removed partially on rota evaporator followed by the dilution with MeOH. The reaction mixture was then column purified (eluted in 75% EtOAc in petroleum ether)



to yield 65% of **7** as white solid. ^1H NMR (200 MHz, CDCl_3): δ 1.83 (s, 3H), 2.06 (s, DMA, COCH_3), 2.91(s, 3H), 2.99(s, 3H), 3.31 (s, 3H), 3.47 (m, 2H), 3.65–4.12 (m, 10H), 4.29 (m, 1H), 5.06 (s, 2H), 5.77 (d, $J = 3.22$ Hz, 1H), 7.30 (m, 5H), 7.64 (s, 1H), 9.97 (bs, 1H); ^{13}C NMR (50 MHz, CDCl_3): δ 12.1, 21.2 (DMA, $-\text{CH}_3$), 29.4, 35.0, 37.8(DMA, $-\text{N}-\text{CH}_3$), 49.9, 58.7, 60.4, 66.5, 68.3, 69.7, 71.1, 77.2, 82.1, 84.4, 88.1, 110.4, 127.7, 127.8, 128.2, 136.1, 136.8, 150.5, 156.3, 164.3, 171.0 (DMA, $-\text{CO}$); **HRMS (EI)**: Mass calculated for $\text{C}_{22}\text{H}_{29}\text{N}_3\text{O}_9\text{Na}$ (M+Na) 502.1796, found 502.1800.

5'-O-Dimethoxytrityl-2'-O-[N-benzyloxycarbonyl-2'-O-R-(2-amino-3-methoxypropyl)] ribothymidine, **31**:

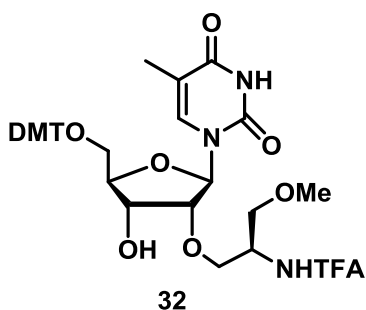
To 2'-O-functionalized ribothymidine **30** (3.0 mmol, 1.5 g) dissolved in dry pyridine (10 mL), was added DMT-Cl (3.3 mmol, 1.1 g) and catalytic amount of DMAP (~20 mg). The reaction mixture was stirred at room temperature for 5–6 h. Pyridine was removed under reduced pressure and the residue dissolved in EtOAc. 10% aqueous NaHCO_3 , water and brine washes were given to the organic layer. The organic layer was dried over anhydrous Na_2SO_4 and concentrated to dryness. The crude compound was column purified



(eluted in 60% EtOAc in petroleum ether) to furnish **31** in 81% yield as a white foam. ^1H NMR (200 MHz, CDCl_3): δ 1.36 (s, 3H), 3.34 (s, 3H), 3.37–3.58 (m, 5H), 3.78 (s, 6H), 3.97–4.09 (4H), 4.41 (m, 2H), 5.09 (s, 2H), 6.81 (m, 2H), 7.28–7.65 (m, 12H), 7.65 (s, 1H), 9.51 (bs, 1H); ^{13}C NMR (50 MHz, CDCl_3): δ 11.7, 29.6, 50.1, 55.1, 59.0, 62.1, 66.8, 68.9, 70.7, 71.5, 77.2, 82.9, 83.5, 86.8, 87.1, 111.0, 113.2, 127.0, 127.9, 128.1, 128.4, 130.0, 135.2, 136.2, 144.2, 150.4, 156.2, 158.6, 164.0; HRMS (EI): Mass calculated for $\text{C}_{43}\text{H}_{47}\text{N}_3\text{O}_{11}\text{Na}$ ($\text{M}+\text{Na}$) 804.3103, found 804.3105.

5'-O-Dimethoxytrityl-2'-O-(N-trifluoroacetyl-2'-O-R-(2-amino-3-methoxypropyl)) ribothymidine, 32:

The 5'-DMT protected 2'-O-functionalized thymidine derivative **31** (2.6 mmol, 2.1 g) was dissolved in MeOH (15 mL) and 10% Pd-C (10% w/w, 0.21 g) was added. Then the reaction mixture was subjected to catalytic hydrogenation at 65 psi of hydrogen pressure for 5 h. After TLC analysis, Pd-C was removed by filtration over celite and concentration of the filtrate *in vacuo* gave free amine. Without further purification, the amine was subjected to trifluoroacetyl protection. The crude amine (2.5 mmol, 1.6

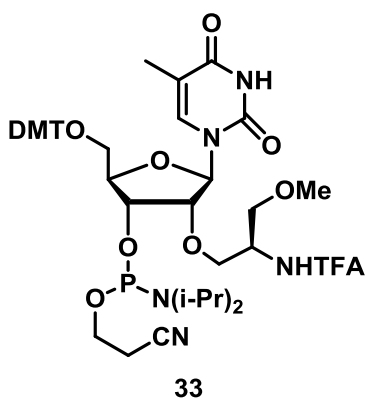


g) was dissolved in MeOH (15 mL), followed by the addition of NEt_3 (3.7 mmol, 0.54 mL). Ethyltrifluoroacetate (xxx mmol, xxxx mL) was added and the reaction mixture stirred at room temperature for 10 h. MeOH was removed on rota evaporator and the reaction mixture was diluted with EtOAc. To the EtOAc containing reaction mixture, water wash and 5% aq. NaHCO_3 wash were given. The organic layer was dried over anhydrous Na_2SO_4 and concentrated *in vacuo*. The crude compound was column purified (eluted in 60% EtOAc in petroleum ether) to yield **32** in 78% over two steps as a white foam. ^1H NMR (200 MHz, CDCl_3): δ 1.36 (s, 3H), 3.38 (s, 3H), 3.45–3.54 (m, 2H), 3.59 (d, $J = 4.55$ Hz), 3.79 (s, 6H), 3.84–3.89 (m, 1H), 4.00–4.11 (m, 3H), 4.31–4.49 (m, 2H), 5.95 (d, $J = 2.91$ Hz), 6.81 (m, 4H), 7.27–7.51 (m, 9H), 7.69 (s, 1H), 9.44 (bs, 1H); ^{13}C NMR (50 MHz, CDCl_3): δ 11.7, 49.3, 55.2, 59.1, 61.9, 68.8, 69.5, 70.5, 83.1, 83.5, 86.8, 87.3, 111.4, 113.1,

113.2, 127.1, 127.7, 127.7, 128.0, 129.0, 130.0, 135.0, 135.2, 144.2, 150.9, 158.6, 163.8;
HRMS (EI): Mass calculated for C₃₇H₄₀F₃N₃O₁₀Na (M+Na) 766.2558, found 766.2561.

Phosphoramidite, 33:

To a solution of **32** (1.3 mmol, 1 g) in dry DCM (15 mL) was added DIPEA (5.8 mmol, 1.0 mL). 2-cyanoethyl-*N,N*-diisopropyl-chloro phosphine (1.6 mmol, 0.35 mL) was added to the solution at 0 °C and the reaction mixture stirred at room temperature for 3 h. The contents were then diluted with DCM and washed with 5% NaHCO₃ solution. The organic phase was dried over anhydrous Na₂SO₄ and concentrated to foam. The residue was redissolved in DCM and precipitated with n-hexane to yield corresponding phosphoramidite **33** in 82% yield as



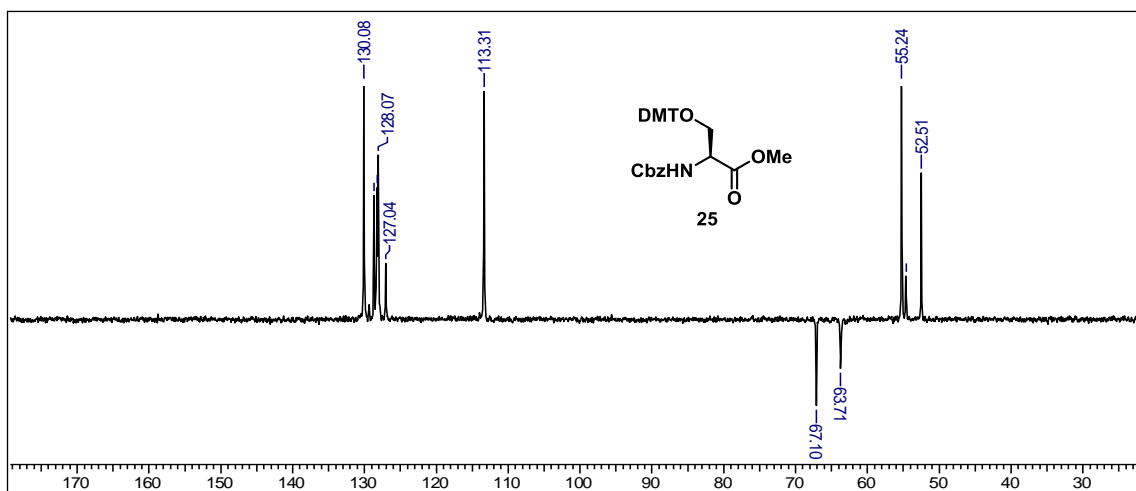
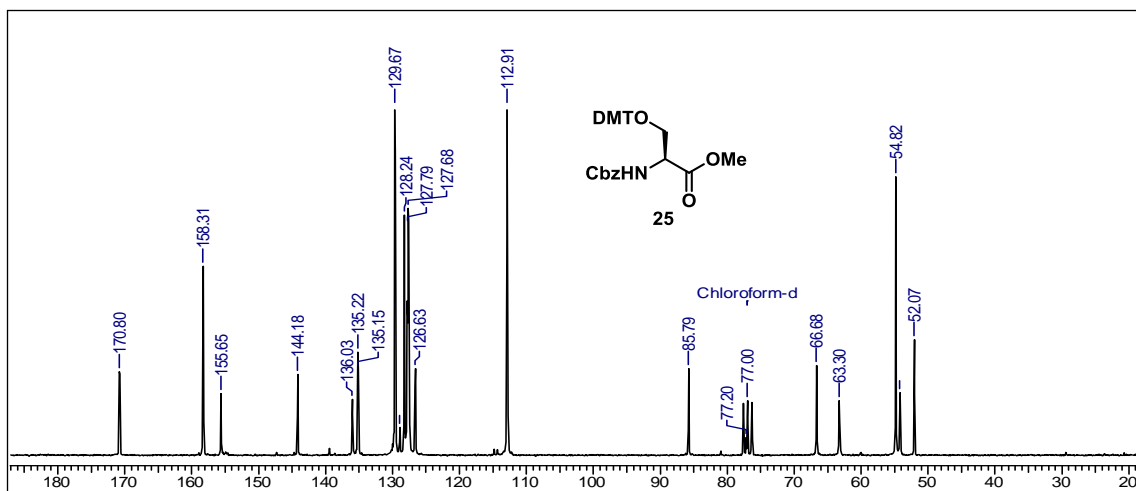
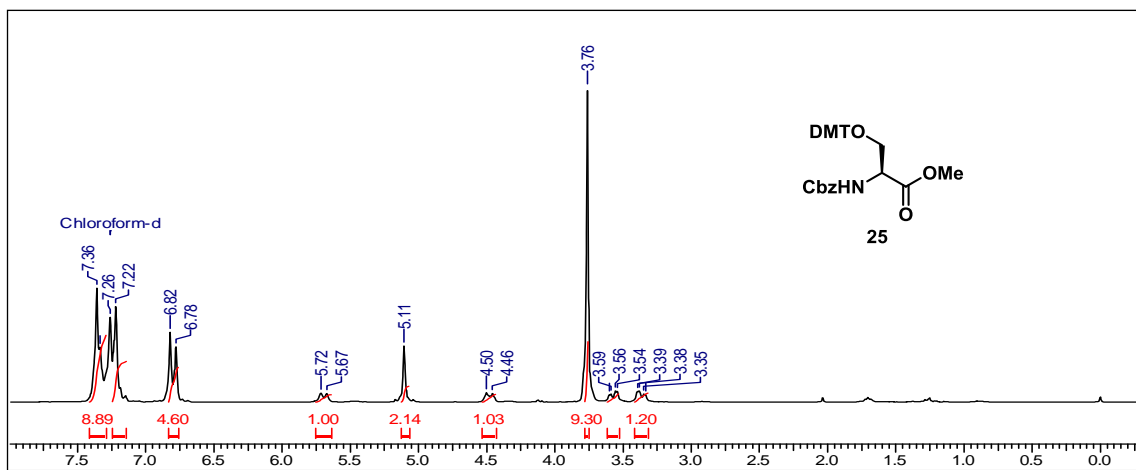
white foam. ³¹P NMR (Acetonitrile, D₂O as external standard, 400 MHz): δ 149.22, 149.68. **HRMS (EI)**: Mass calculated for C₄₆H₅₈F₃N₅O₁₁P (M+H) 944.3817, found 944.3824.

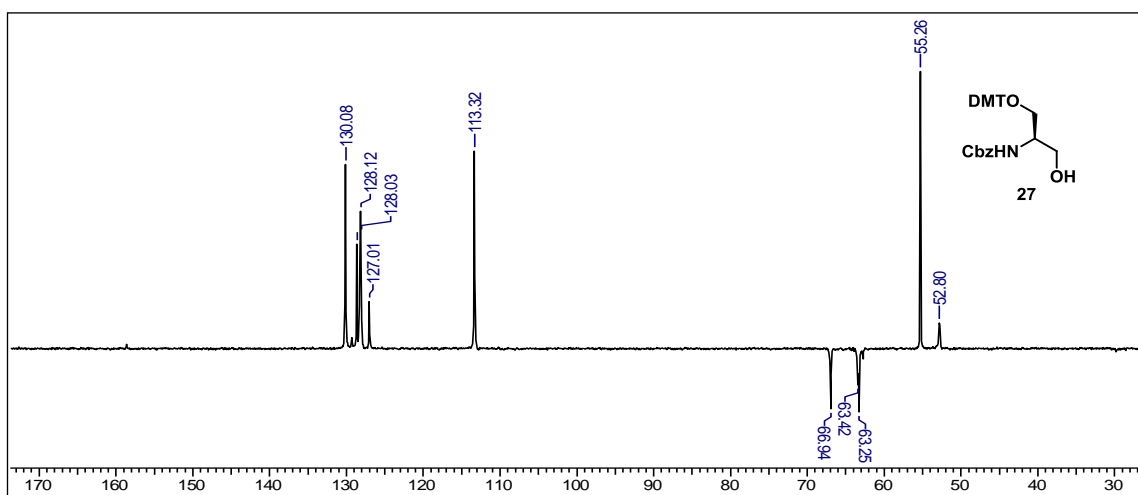
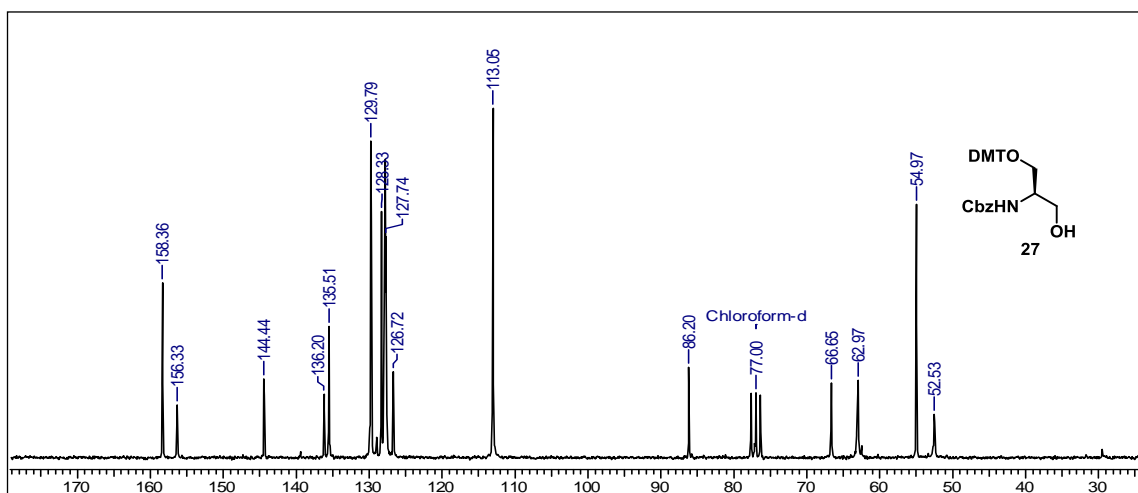
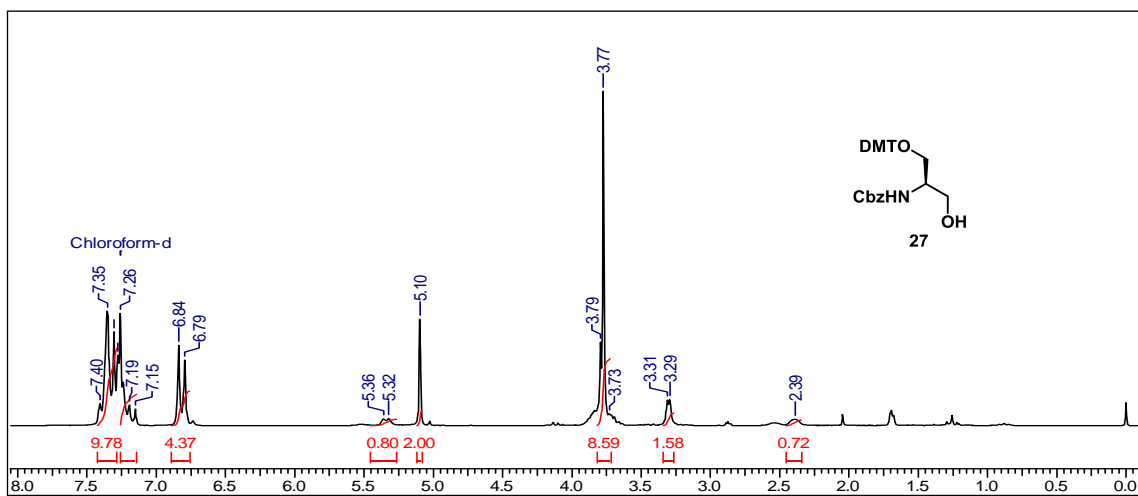
2B.8 Appendix

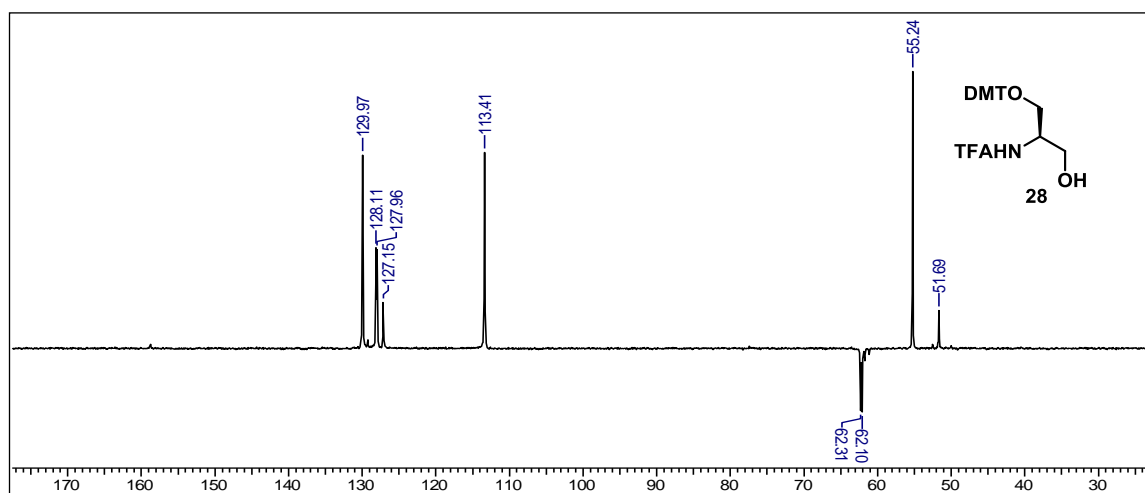
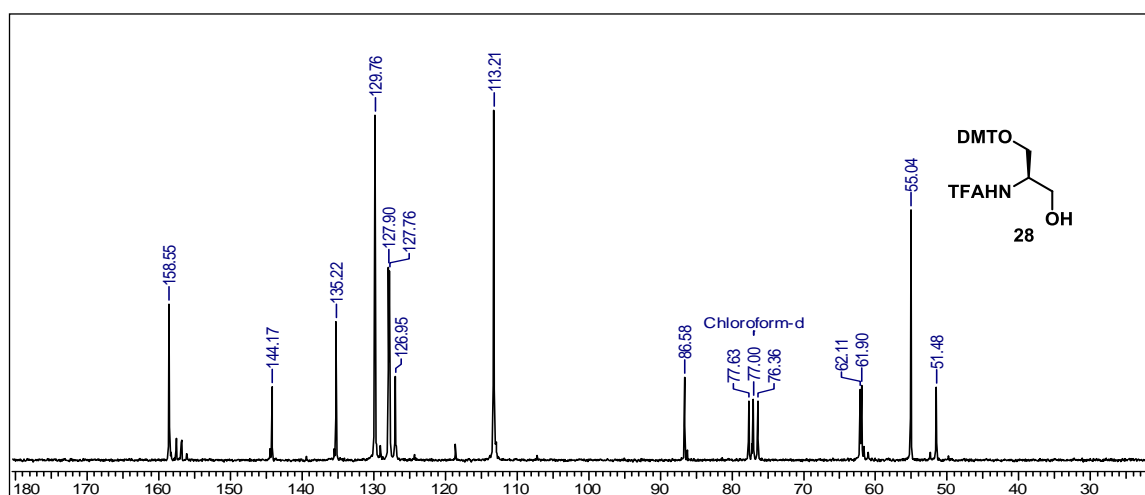
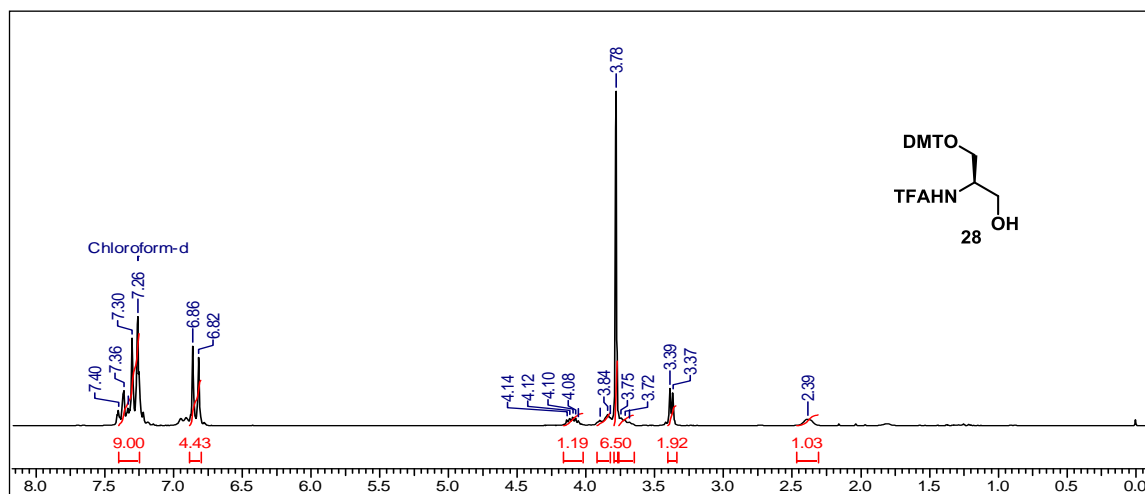
Compounds - Spectral data	Page No.
25- ¹ H, ¹³ C NMR & DEPT	119
27- ¹ H, ¹³ C NMR & DEPT	120
28- ¹ H, ¹³ C NMR & DEPT	121
29- ³¹ P, 30- ¹ H, & ¹³ C NMR	122
30- DEPT, 31 - ¹ H, ¹³ C NMR	123
31- DEPT, 32 - ¹ H, ¹³ C NMR	124
32- DEPT, 33- ³¹ P NMR	125
27, 28- HRMS	126
29, 9- HRMS	127
10, 30- HRMS	128
31, 32- HRMS	129
33- HRMS	130
ON15, ON16- MALDI-TOF spectra	131
ON17, ON18- MALDI-TOF spectra	132
ON19, ON20- MALDI-TOF spectra	133

Chapter 2

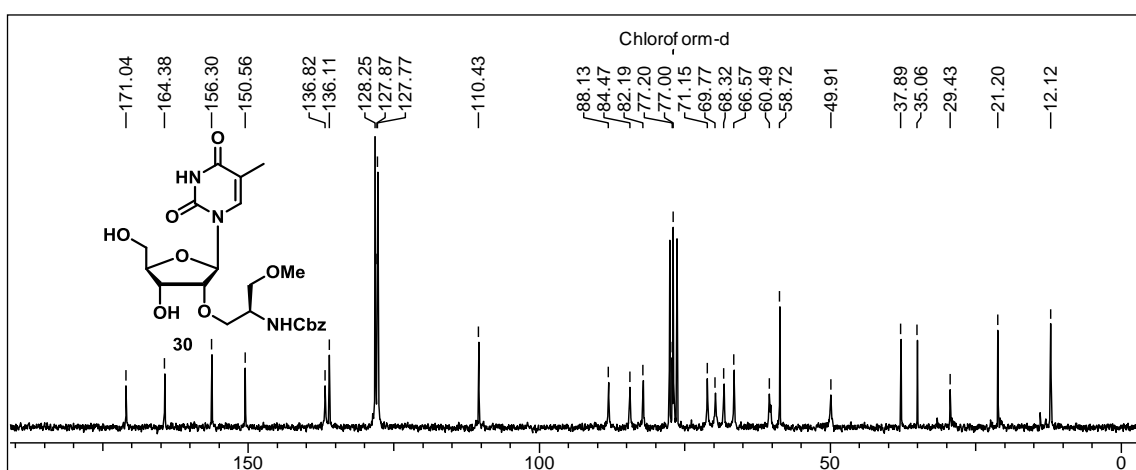
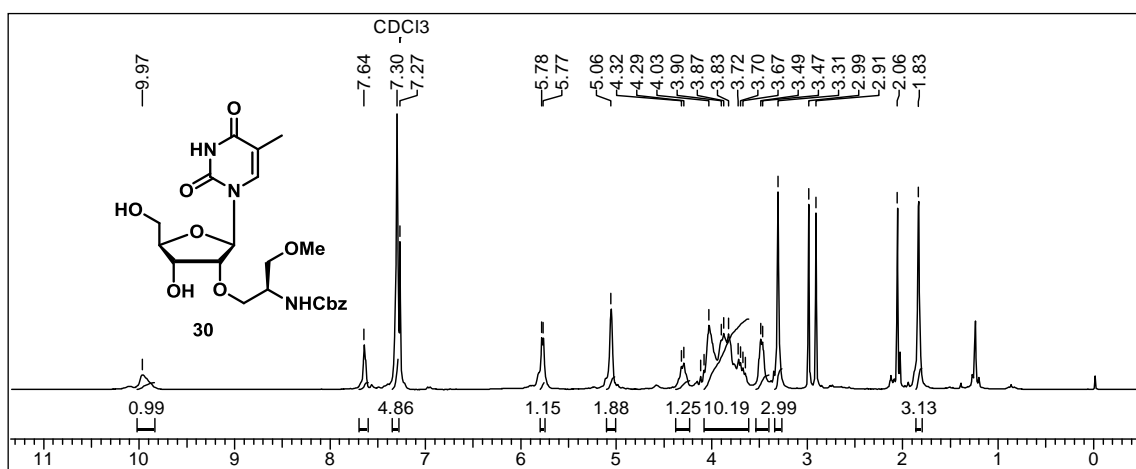
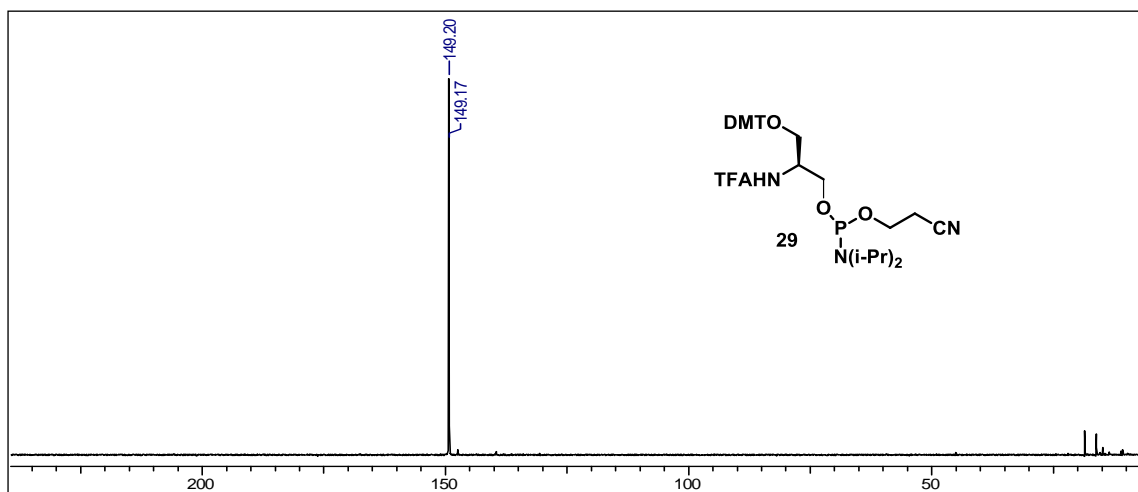
NMR spectral data:

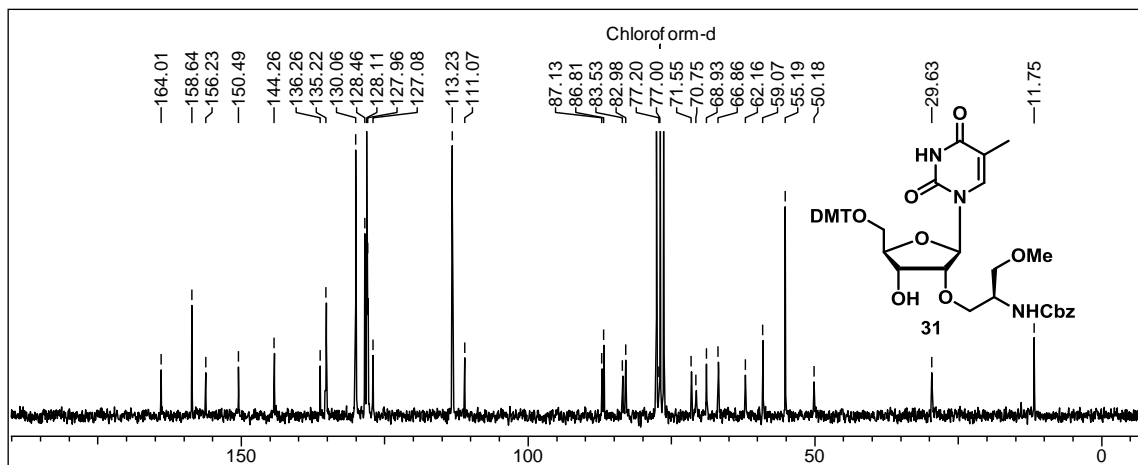
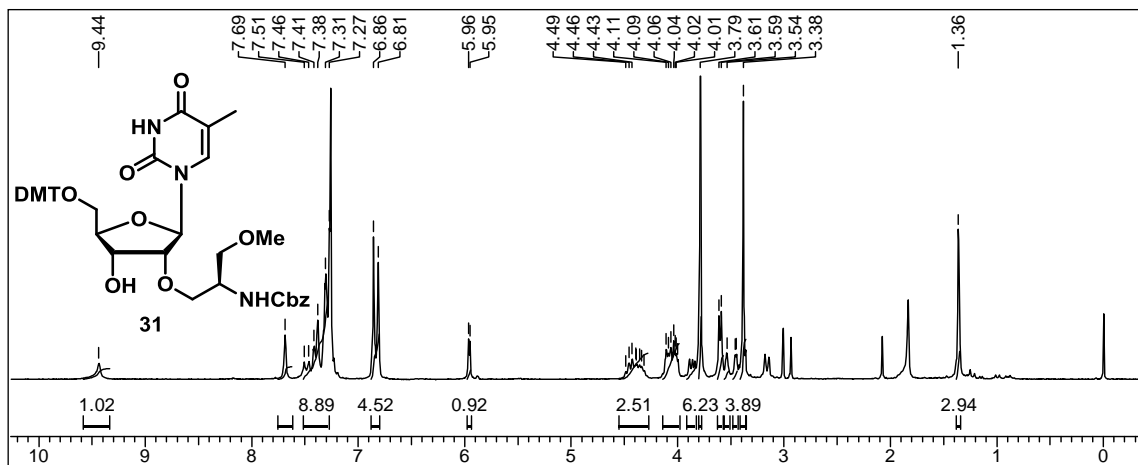
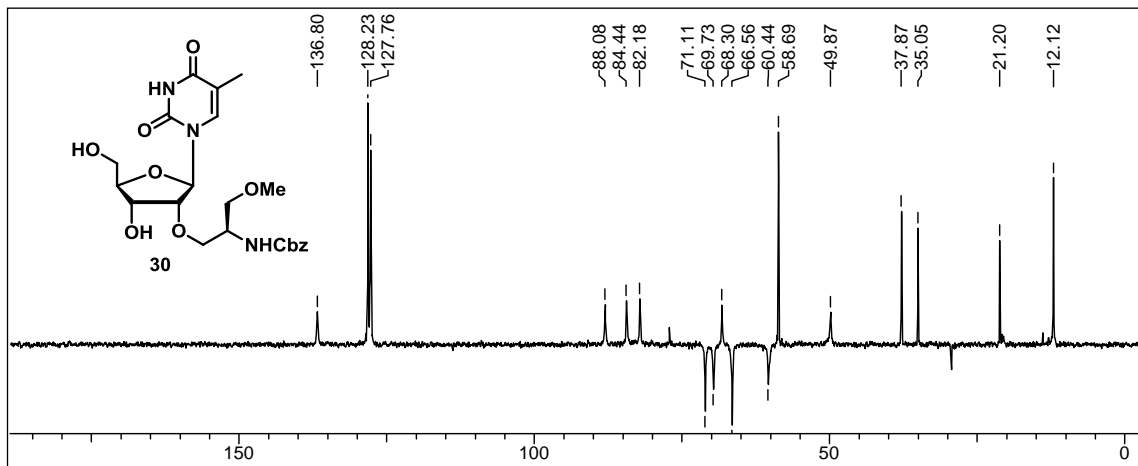


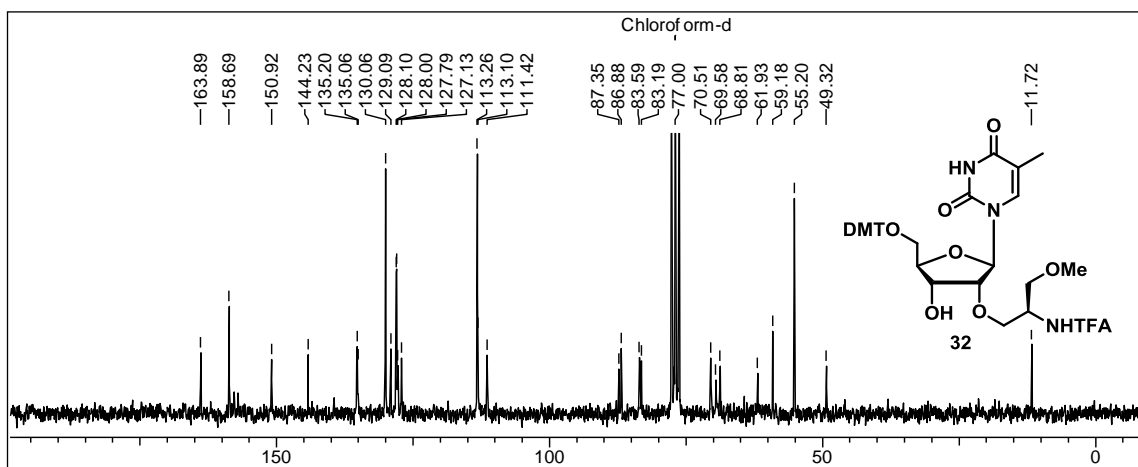
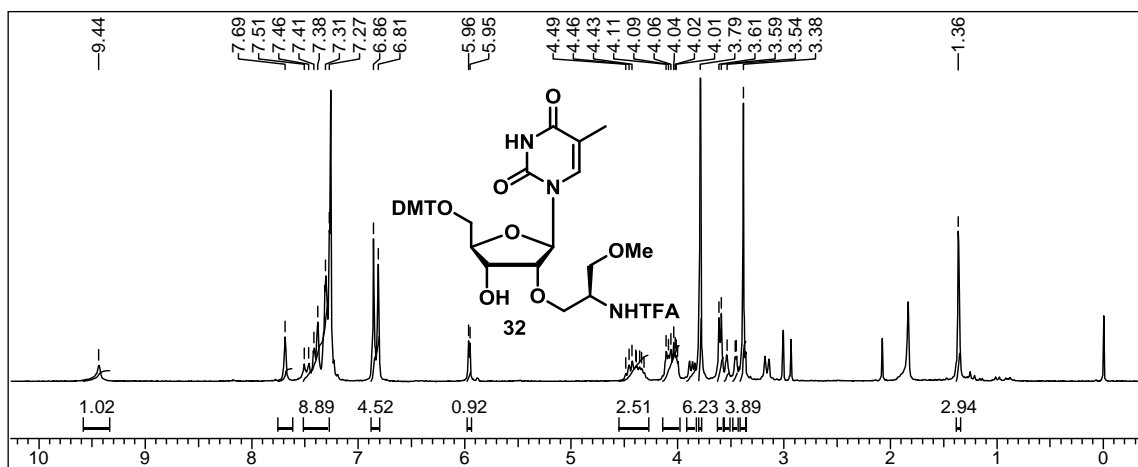
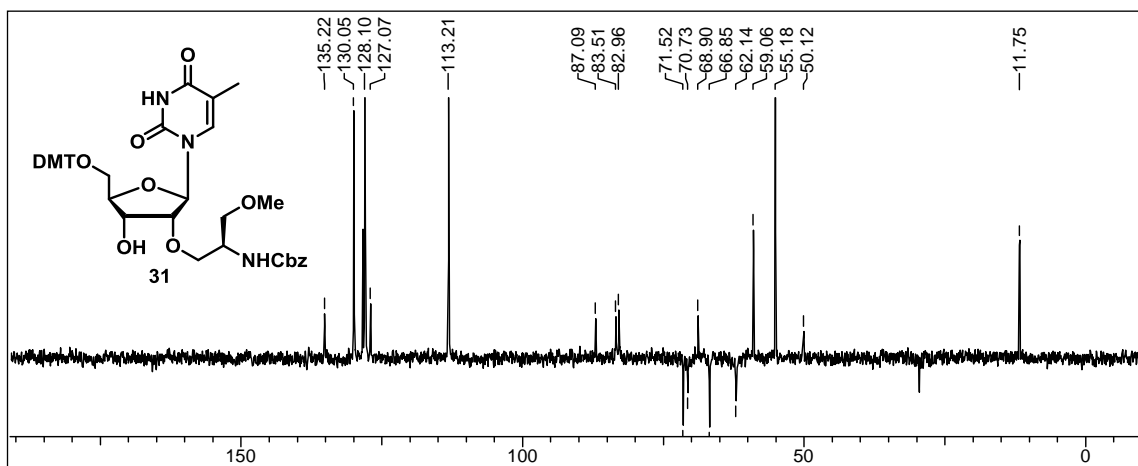


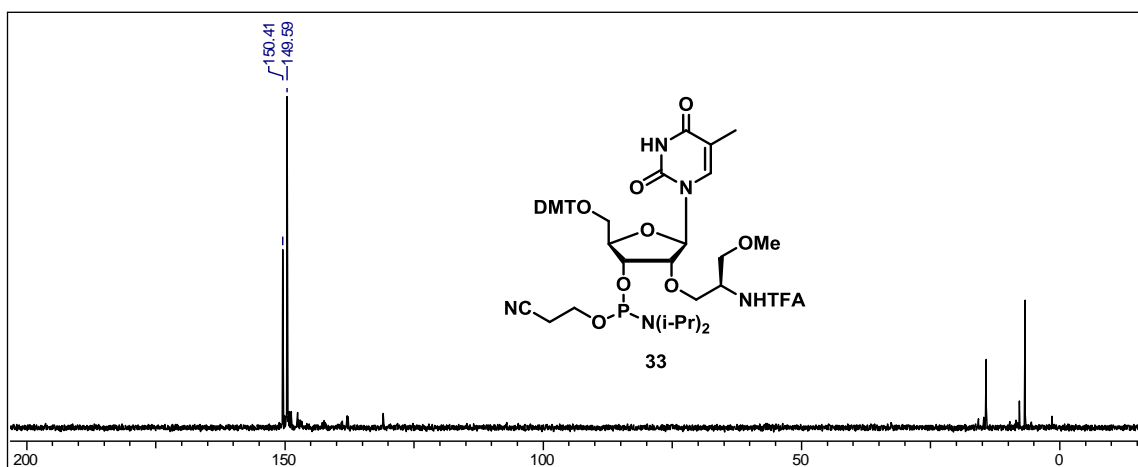
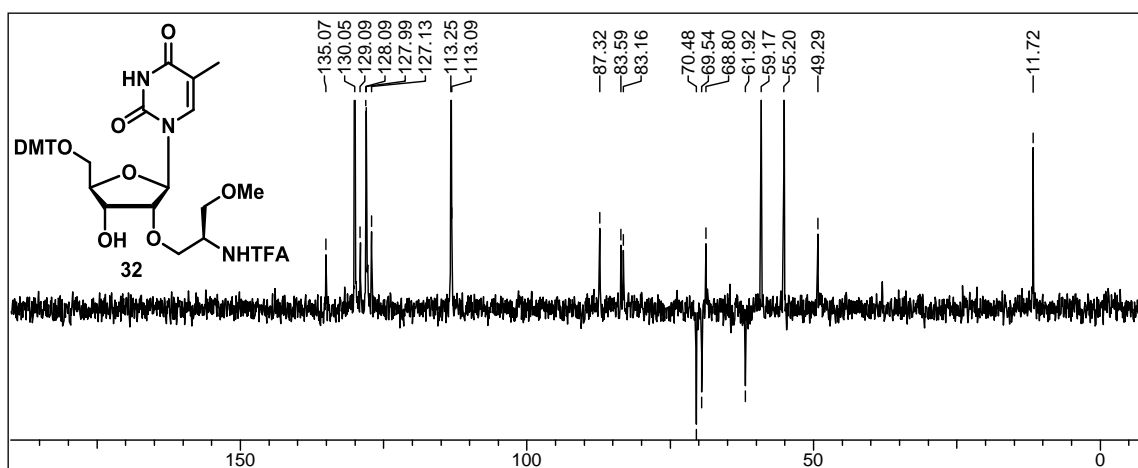


Chapter 2



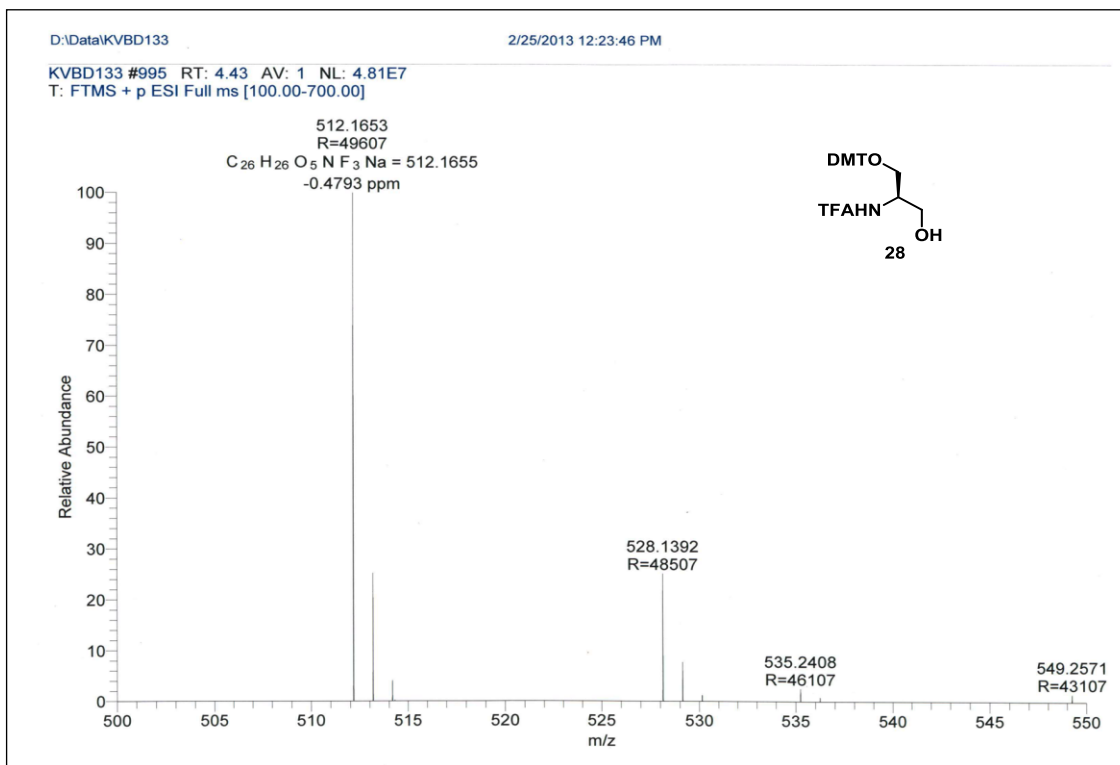
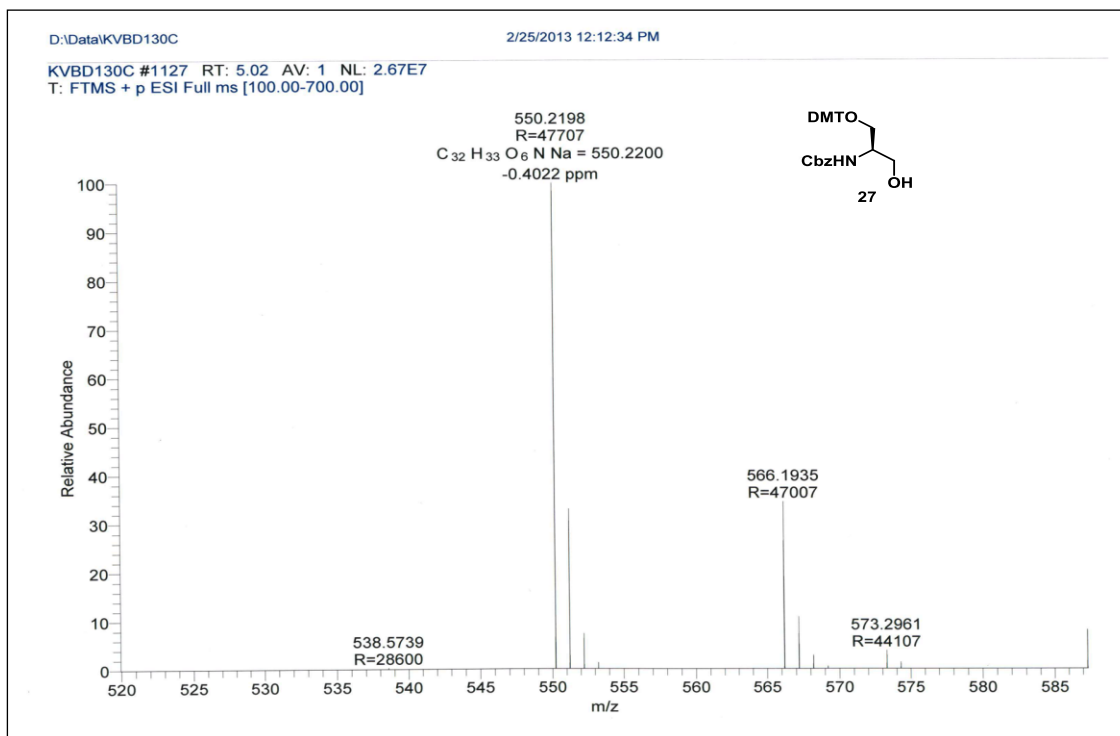




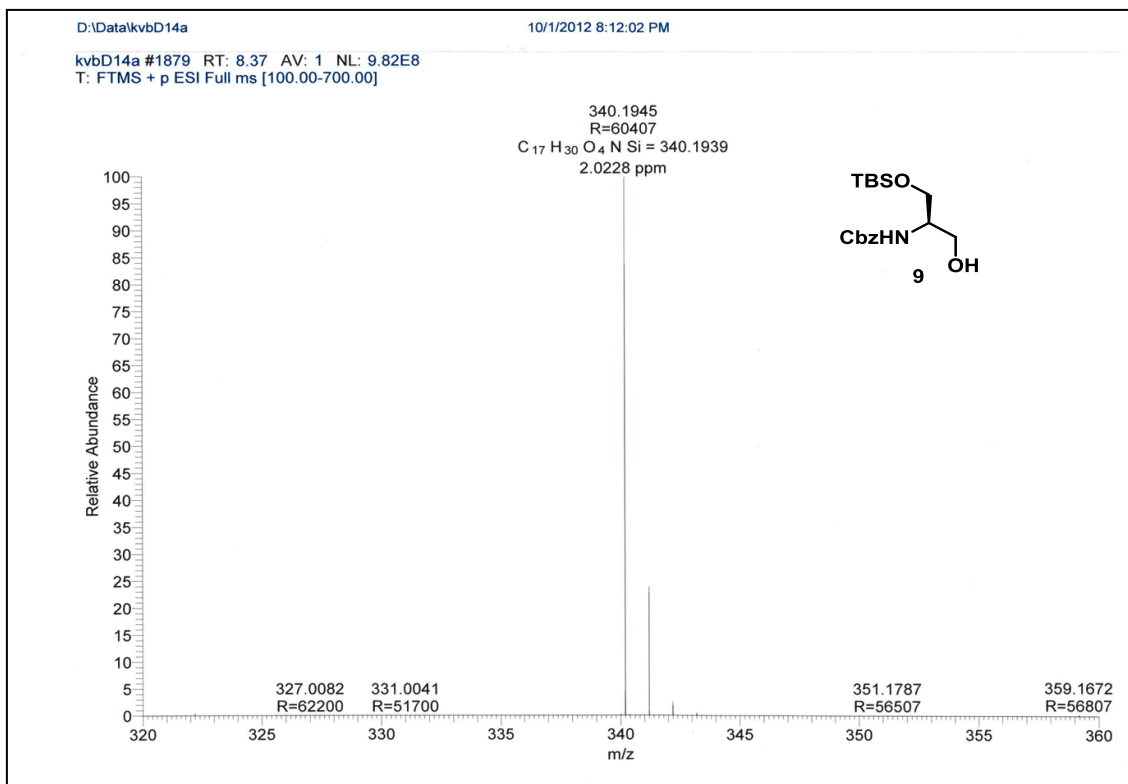
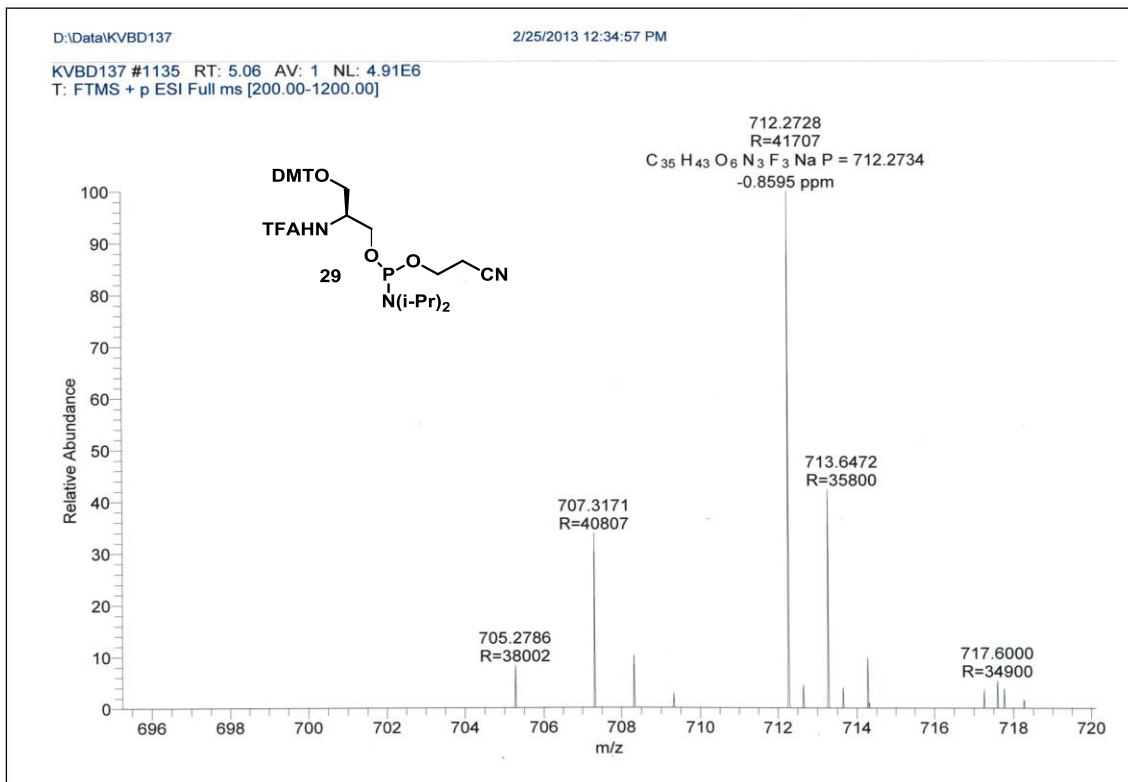


Chapter 2

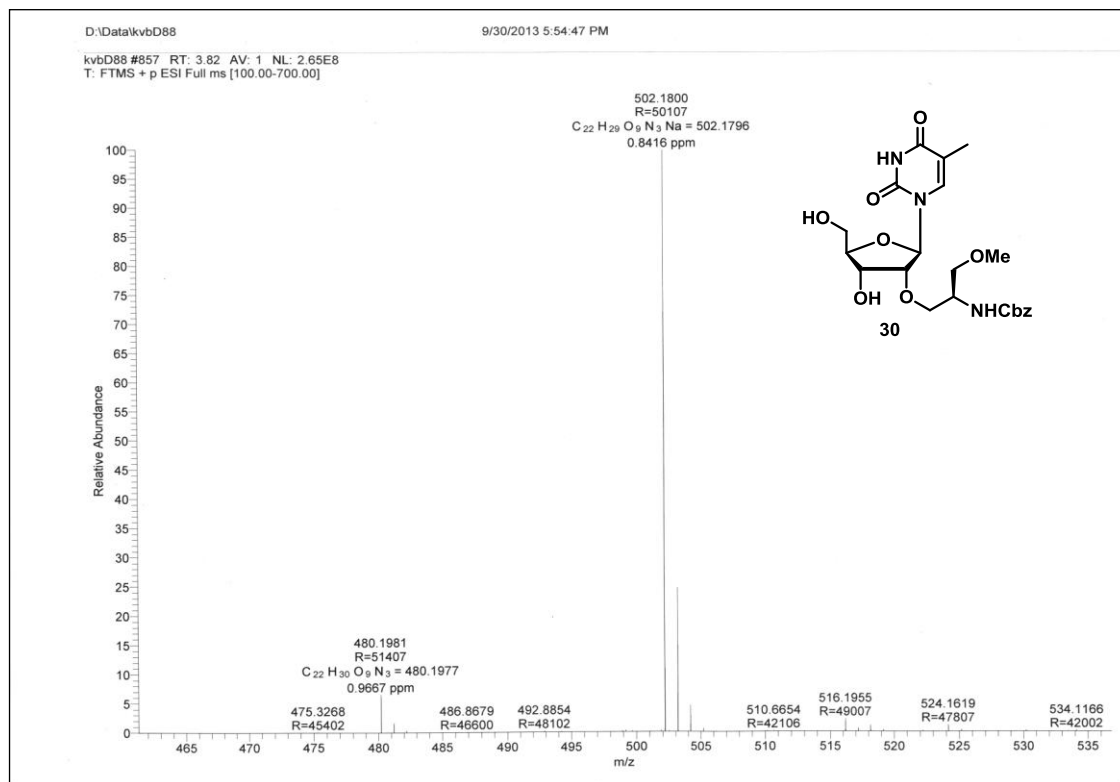
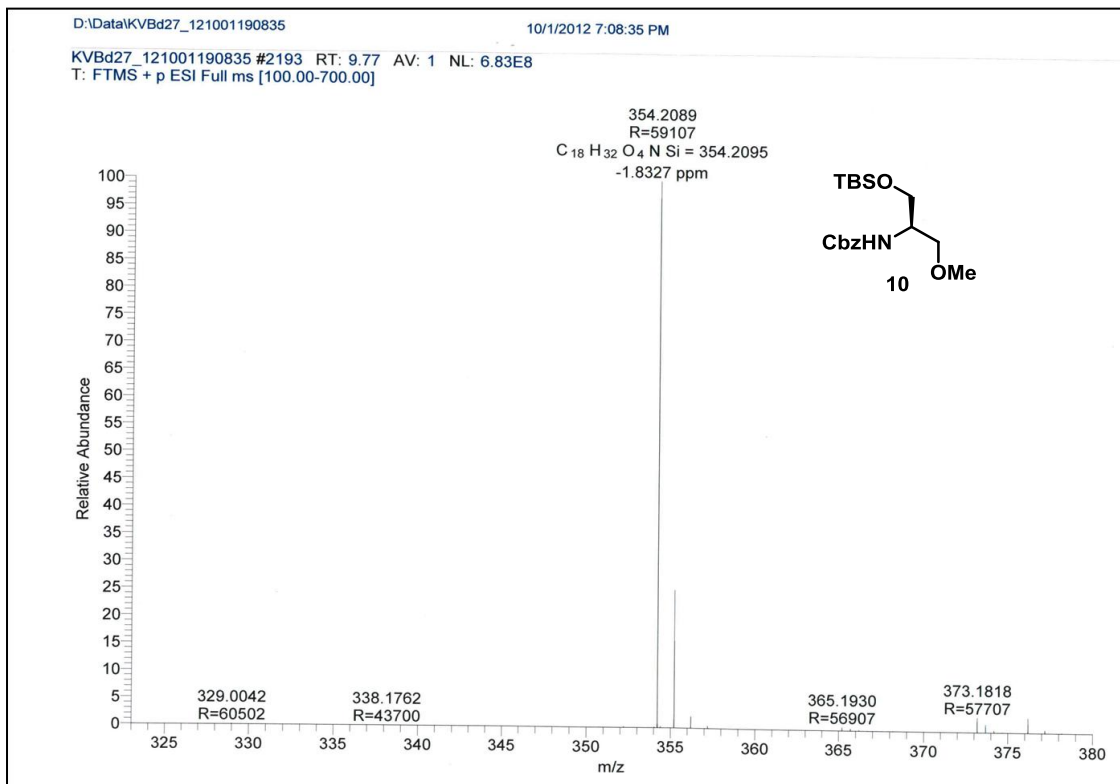
HRMS Spectral data:



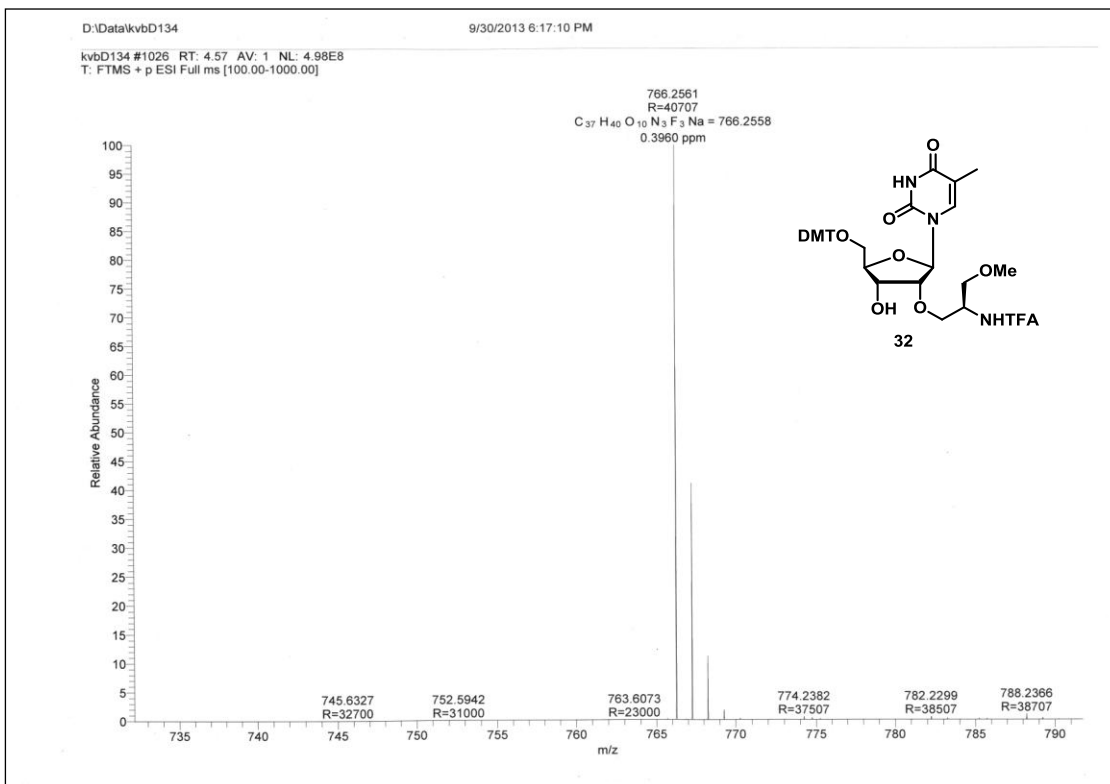
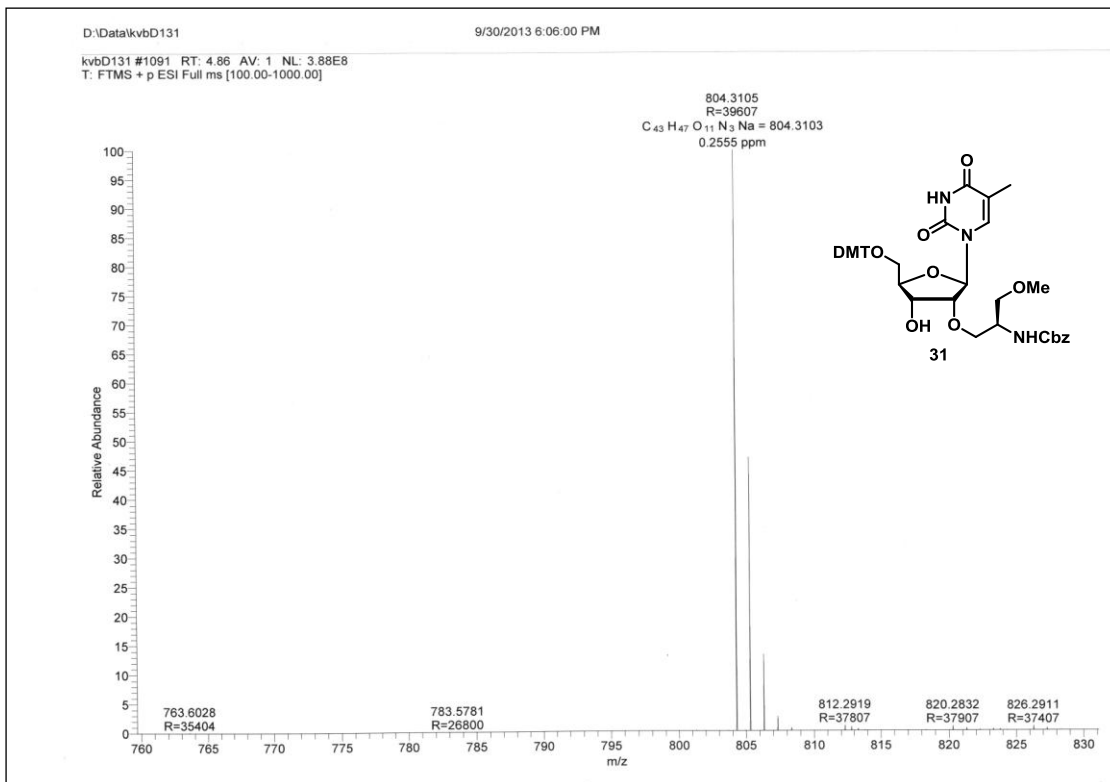
Chapter 2



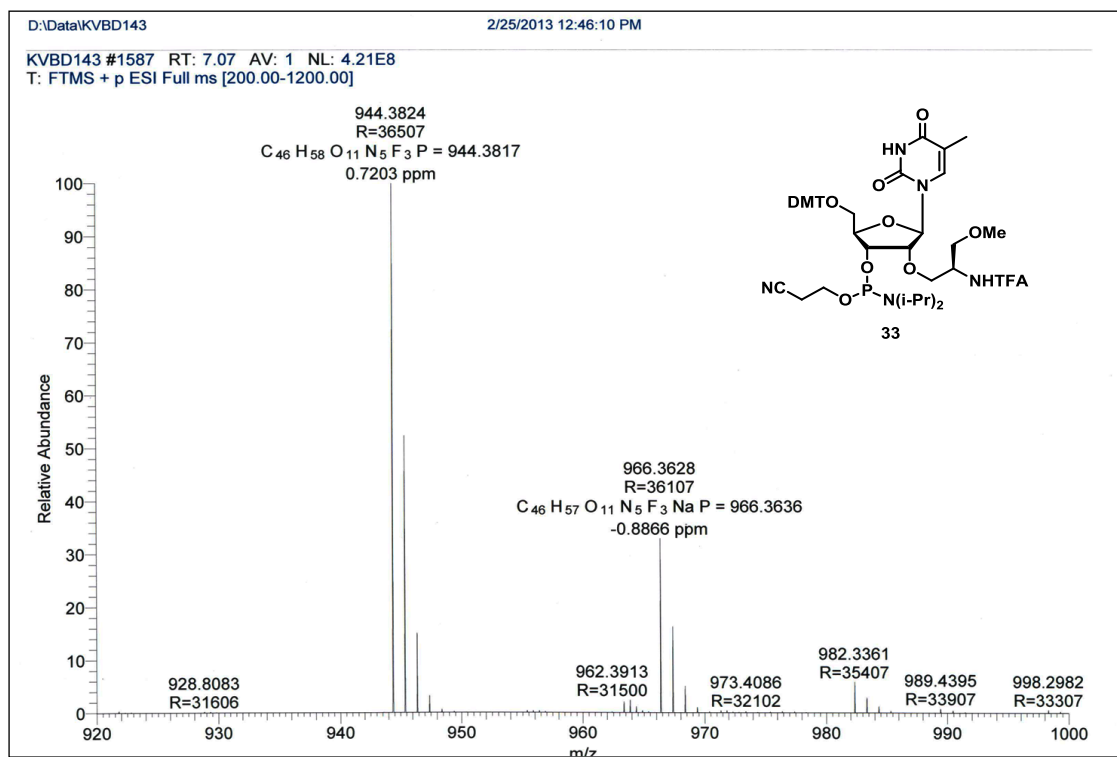
Chapter 2



Chapter 2

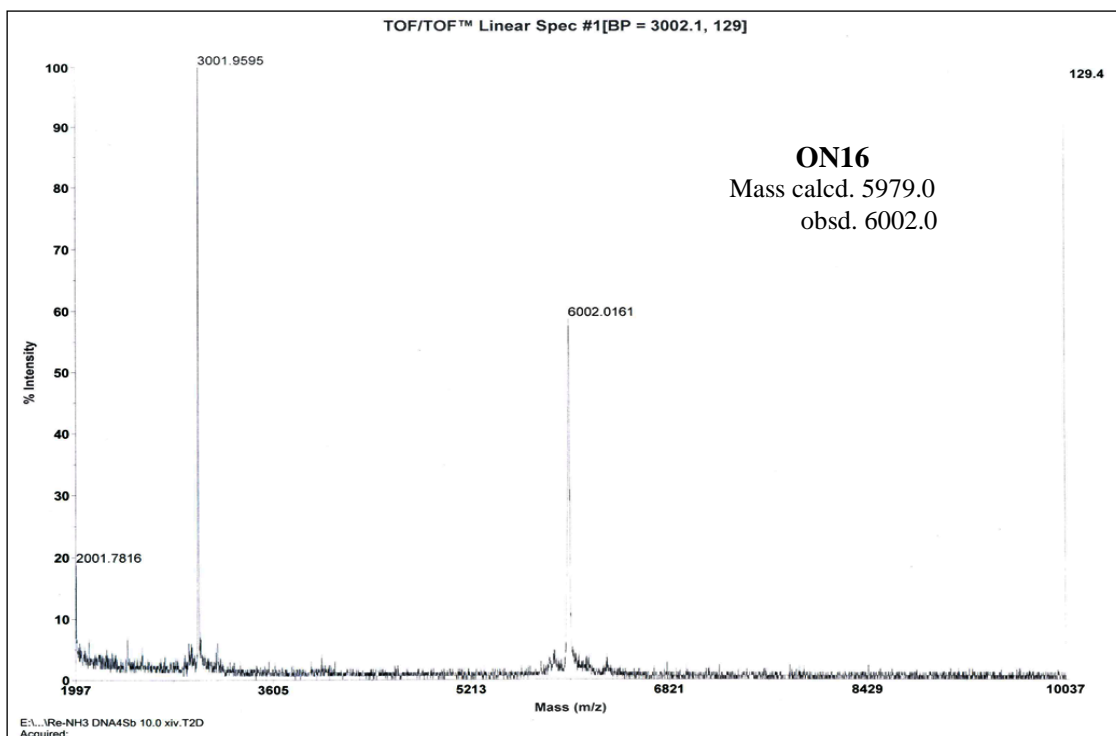
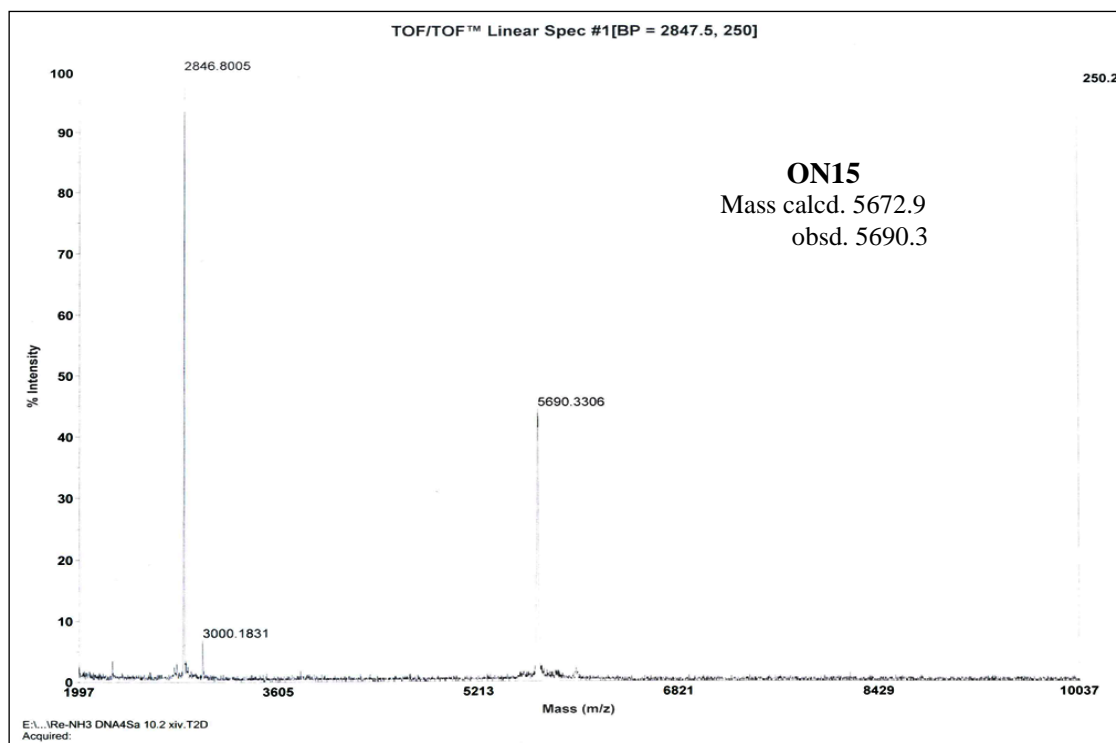


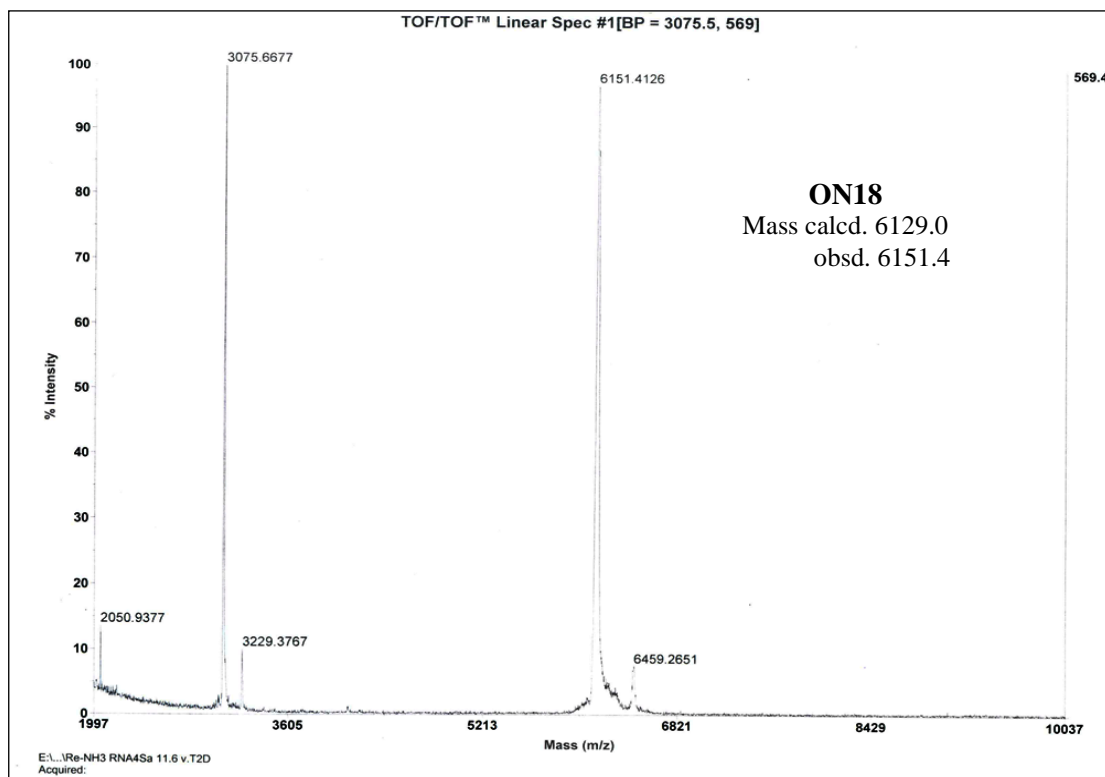
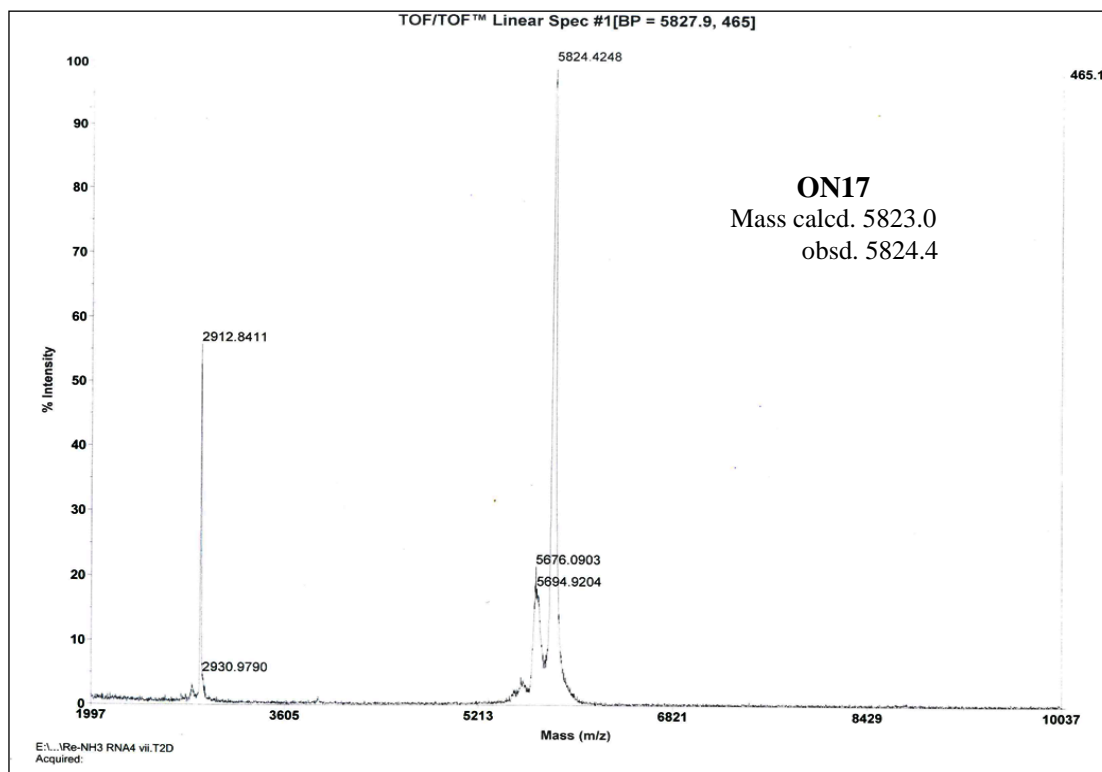
Chapter 2

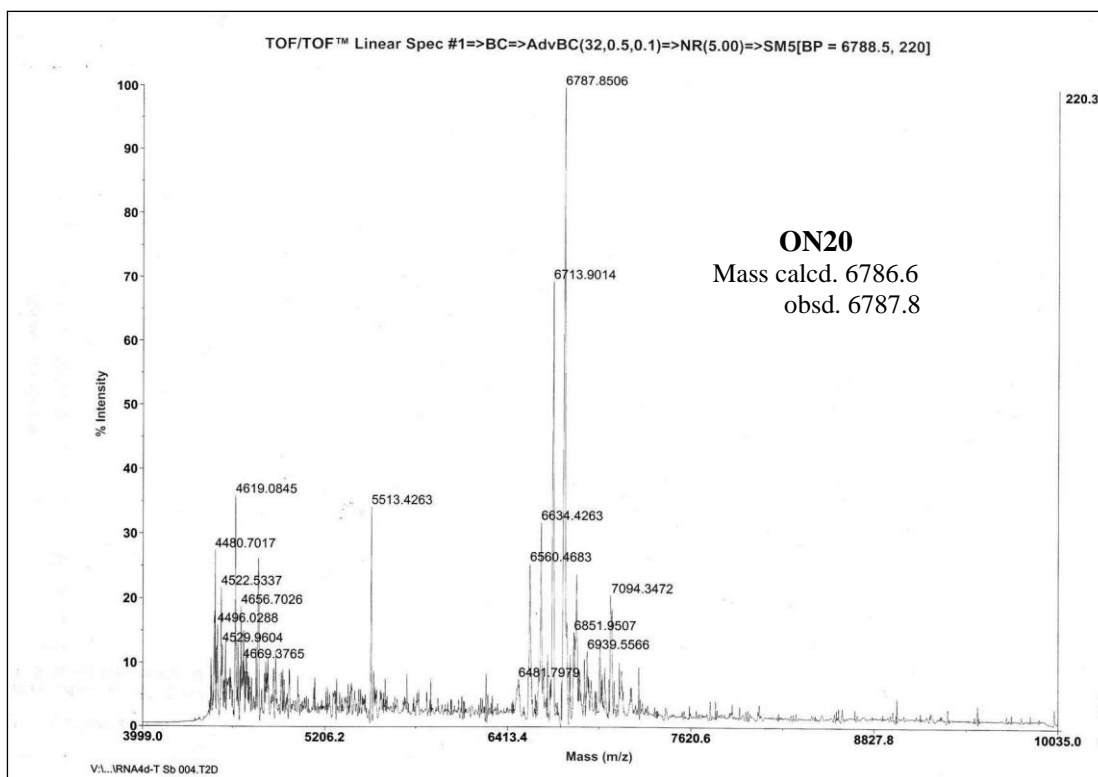
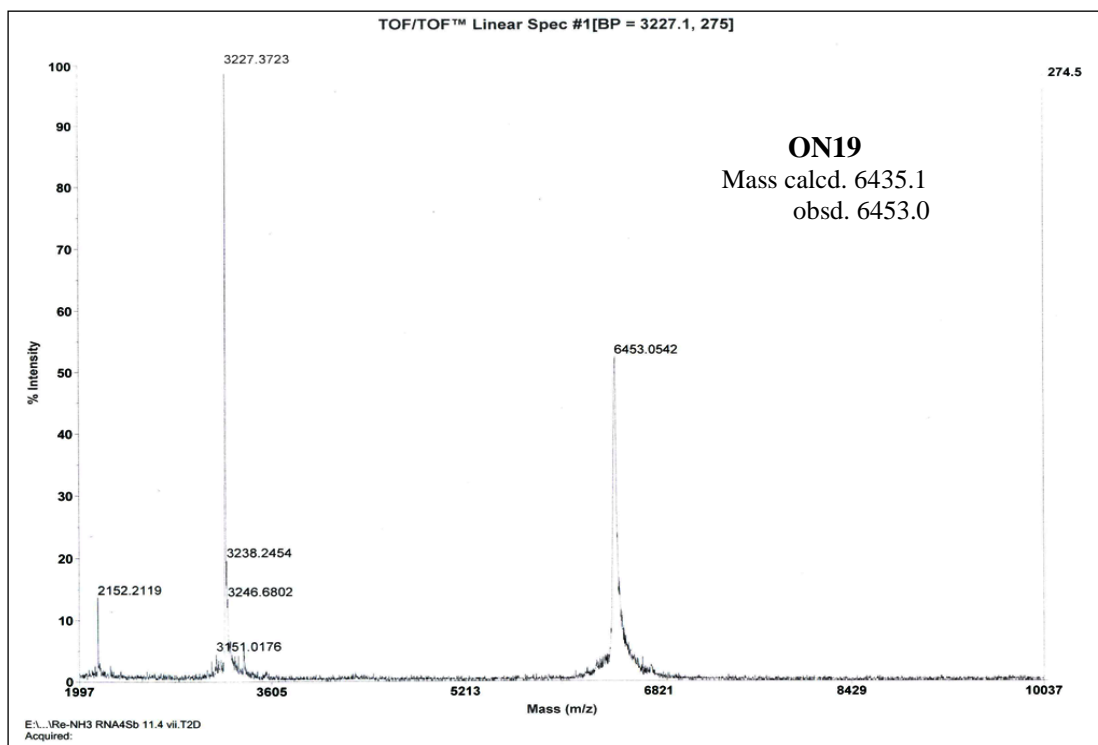


Chapter 2

MALDI Spectral data:







Section 2C: Design, synthesis and biophysical evaluation of 2'-guanidino functionalized oligonucleotides

2C.1 Introduction

The quest for potential antisense drugs has witnessed the development of several chemically modified oligonucleotides. As discussed in the previous sections, the important parameters to be considered during the design of chemical modifications for a better antisense therapy drug are to render highly specific and strong binding capability to the target RNA/DNA as well as stability to the enzymes (nucleases, proteases, etc.) and efficient cellular uptake. Many of the chemical modifications have satisfactorily addressed the high binding affinity to the target and nuclease stability, but little has been achieved in enhancing the cellular uptake.⁵⁶

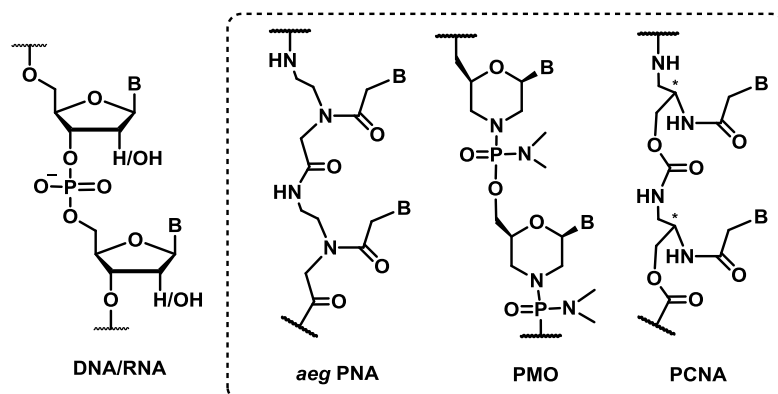


Figure 15 Neutral backbone-oligonucleotides to reduce the charge repulsions while targeting DNA/RNA.

Nucleic acids are well known to form duplexes and triplexes through Watson-Crick and Hoogsteen base-pairing. During duplex/triplex formation, their negatively charged backbones certainly suffer the inter strand repulsions, which would have considerable effect on the stability of the duplex/triplex. For therapeutic applications to target any DNA/RNA with enhanced affinity, an obvious/easy option is to reduce the inter-strand repulsions. There are two ways of achieving this: one is the complete replacement of charged phosphate backbone by a neutral backbone like peptide, carbamate, etc., (Figure 15) and the other one is to reduce the net anionic charge by introducing functional groups that are

Chapter 2

cationic in nature at physiological pH (Figure 16). Accordingly, reported oligonucleotides containing neutral backbones, such as peptide nucleic acids (PNAs),⁵⁷ or phosphorodiamidate morpholino oligomers (PMOs)³⁴ and polycarbamate nucleic acids (PCNAs)⁵⁸ bind their target with higher affinity than their negatively charged counterparts without loss of specificity (Figure 15).

The second interesting approach is to neutralize the inter-strand repulsion by introducing the cationic functional group links to the nucleotide through the phosphate backbone,⁵⁹ 2'-*O*-functionalization⁶⁰ or nucleobase.⁶¹ These zwitter-ionic modified oligonucleotides showed properties of enhanced binding affinities to complementary strands and also high nuclease resistance. Earlier studies focused on the cationic linkages that mostly comprise amine derivatives prone to protonation at physiological pH, this trend being prevalently dominated by the functional group “guanidine” (pK_a 12.5), a functional group that is positively charged over a wide pH range and profoundly credited for efficient cellular uptake.⁶² An interesting example is deoxyribose nucleic guanidine (DNG), where the internucleosidic negatively charged phosphate linkage was replaced by the guanidinium cationic linkage, which resulted in stronger binding to complementary DNA and RNA than their counterparts at physiological ionic strength.⁶³ The enhanced stability by electrostatic interactions was supported by experiments at varying ionic environments.

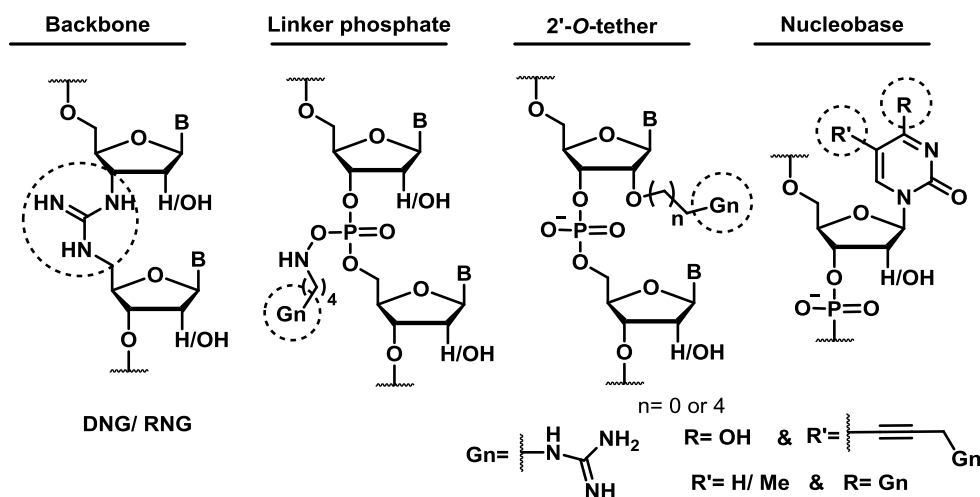


Figure 16 Ionic modifications of DNA/RNA with varied constitution of guanidine group.

Though PNAs are credited as one of the most potential modification for antisense therapeutics, their poor cellular uptake hampered further applications in therapeutics. This major drawback of PNAs was addressed by introducing a guanidinium group through a backbone linker, derived from arginine.⁶⁴ Also, introduction of guanidines in PNAs or LNAs (locked nucleic acids) further enhanced the duplex stability, as exemplified by Gu-petPNA⁶⁵ or GuLNA.⁶⁶ Thus, the introduction of guanidine in these promising modifications enhanced their superiority in the development of antisense therapeutics.

2C.2 Objective and rationale behind the design

The basic chemical difference between DNA and RNA is absence and presence of the 2'-OH group respectively. Hence, 2'-substitution could be the factor governing the sugar pucker ($S \leftrightarrow N$) in DNA/RNA. Intensive research in the area of modified oligonucleotides for a better antisense candidate has resulted in some promising modifications at 2'-position of the sugar ring. The electronegative substituents such as oxygen (as in RNA) or fluorine can shift the ribose conformational equilibrium towards the C3'-*endo* pucker that can form stable duplexes with target RNA. Although the 2'-OMe¹ and 2'-F² modifications formed stable duplexes with the target RNA, they could not confer adequate metabolic stability to the antisense oligonucleotides and hence had to be used with phosphorothioate linkages.^{3b, 39b, 40}

In our design, we substituted the 2'-position of the deoxyribose sugar of DNA with the guanidinium group- (Figure 17), three amines connected to a central carbon in a plane, which remains protonated over a wide pH range. Once protonated, guanidinium can act as an electron withdrawing group and hence may affect the ring puckering of ribose sugar, like 2'-OMe or 2'-F substituted oligonucleotides (Figure 17). The C3'-*endo* conformation of a nucleoside-sugar in AONs having pseudoaxially projected nucleobases is known to stabilize the duplex with target RNA through efficient π -stacking. We anticipated that the 2'- guanidinium group (protonated over $\text{pH} < \sim 12$) would drive the sugar conformation towards C3'-*endo* through additional gauche and strengthened anomeric effects and hence result in increased duplex stability with target RNA. Our design is a new dimension to the

guanidine functional group for directing the conformational equilibrium ($S \leftrightarrow N$). It is also expected to render high nuclease resistance to the AONs along with enhanced binding strength to the target similar to 2'-guanidinoethyl- containing AONs,^{60b} and enhanced cellular uptake like GPNAs.⁶⁴

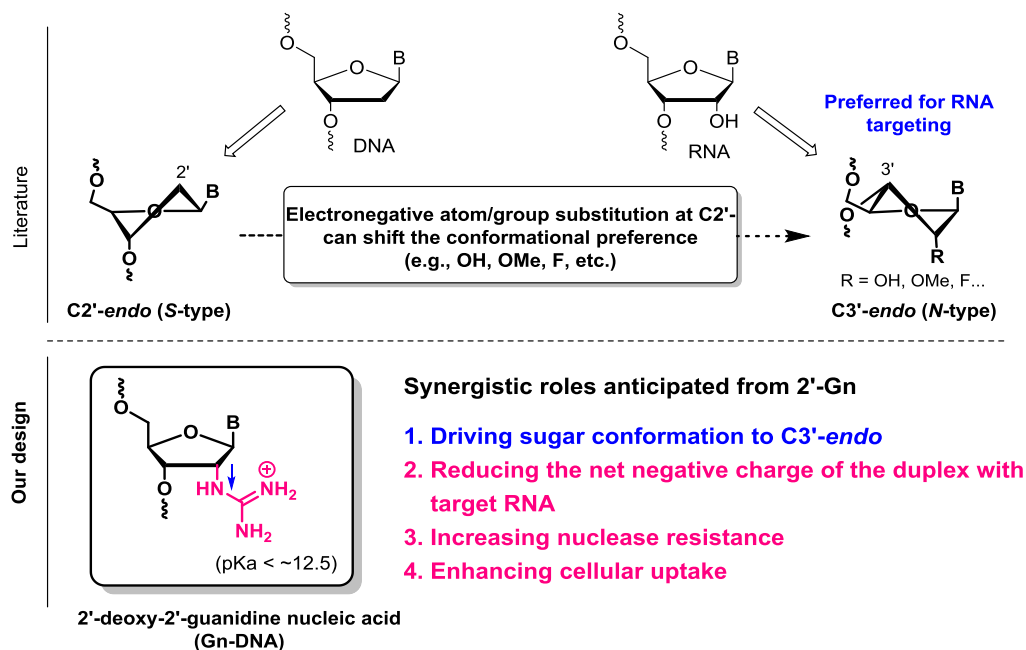


Figure 17 Design of Gn-DNA to play a synergistic role towards potential antisense applications.

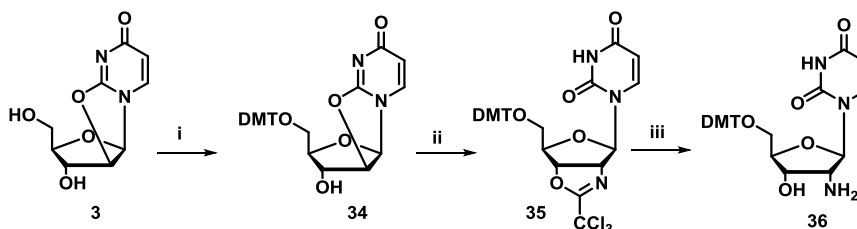
2C.3 Synthesis of 2'-amino-2'-deoxyuridine

Towards the synthesis of 2'-Gn nucleoside, we envisaged guanidinylation of 2'-deoxy-2'-amino-ribonucleoside using commercially available reagents. So, first we undertook the synthesis of 2'-amino-2'-deoxyuridine from uridine. The synthesis of 2'-amino-2'-deoxyuridine can be achieved by the intramolecular cyclization of 3'-tethered trichloroacetimidate followed by basic hydrolysis.⁶⁷

The synthesis began with the protection of the 5'-hydroxyl group of 2,2'-anhydrouridine **3** as its DMT ether **34**. Compound **34** having free 3'-hydroxy is treated with trichloroacetonitrile and triethylamine under reflux conditions to give the oxazoline **35** (Scheme 7). Basic hydrolysis of oxazoline **35** with 6N NaOH in ethanol under reflux

conditions furnished the 2'-deoxy-2'-aminouridine derivative **36**, precursor to the synthesis of 2'-deoxy-2'-guanidinouridine.

Scheme 7 Synthesis of 2'-amino-2'-deoxyuridine



Reagents and conditions: (i) DMT-Cl, pyridine, cat. DMAP (ii) CCl_3CN , TEA, reflux (iii) 6N NaOH/EtOH, 16 h, 60% yield over three steps.

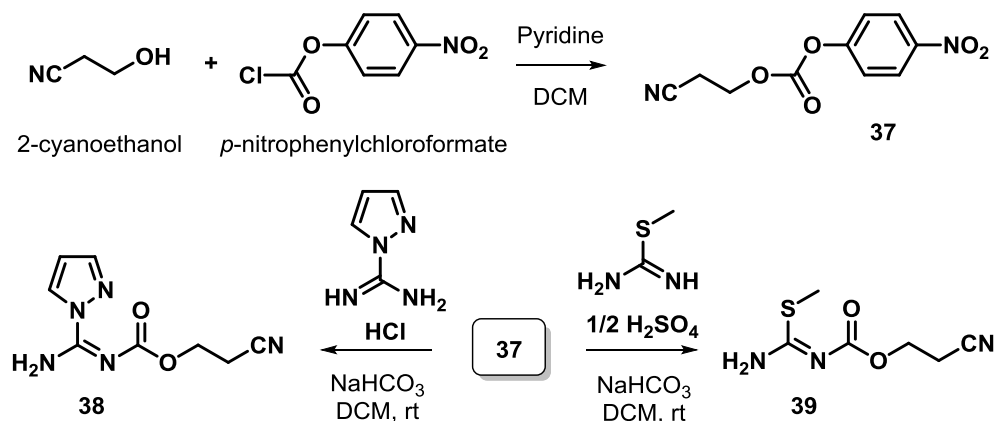
2C.4 Guanidinylation of 2'-deoxy-2'-aminouridine

The guanidinylation of 2'-deoxy-2'-aminouridine derivative **36** proved difficult, as several attempts using commercially available guanidinylation reagents failed. We approached the 2'-Gn-functionalization of uridine in two ways: one was through the nucleophilic attack of the 2'-amine in compound **36** on electrophilic guanidinylation agents and the other was through the $\text{BF}_3 \cdot \text{OEt}_2$ -mediated electrophilic ring opening of 2,2'-anhydrouridine using guanidine as a nucleophile. In addition, the final amidite monomer should contain base-labile protecting groups for guanidine, compatible with automated solid phase synthesis. We chose the recently developed base-labile cyanoethoxycarbonyl- (CEOC) protecting group for the guanidine.^{60b, 68}

We synthesized a new reagent, 2-cyanoethyl(4-nitrophenyl)carbonate **37** for -CEOC protection of the guanidinylation agent and demonstrated the mono -CEOC-protection of 1*H*-pyrazole-1-carboximidine and *S*-methylthiourea (**38** and **39** respectively). The reagent preparation involved the use of relatively cheap *p*-nitrophenylchloroformate for the activation of 2-cyanoethanol rather than disuccinimidyl carbonate.⁶⁹ However, the 2'-guanidinylation of compound **36** was unsuccessful using mono-CEOC-protected guanidinylation reagents **38** and **39**. As the protected reagents failed to guanidylate the 2'-amine of compound **36**, we planned to use sterically less crowded protecting group-free guanidinylation agents (Scheme 9, conditions iii and iv). However, even the free

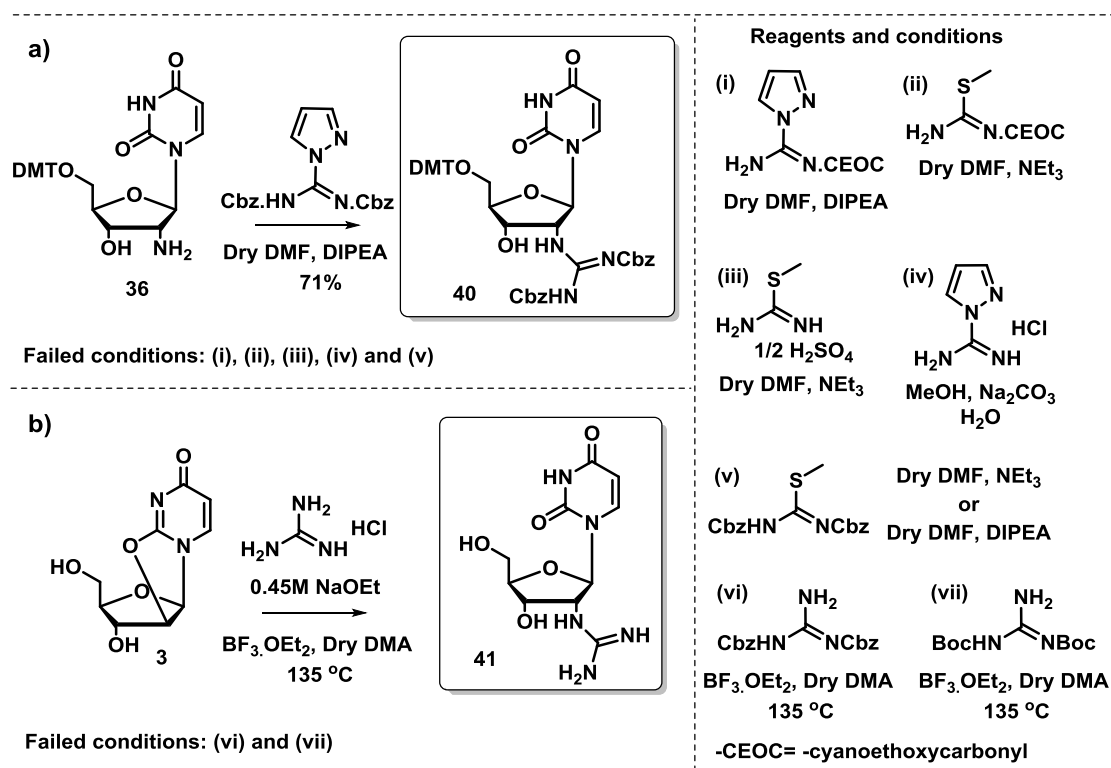
guanidinylation agents 1*H*-pyrazole-1-carboxamide and *S*-methylthiourea did not deliver the 2'-guanidylated product. Guanidinylation attempts using *N,N'*-bis(benzyloxycarbonyl)-*S*-methylthiourea were also unsuccessful. Successful guanidinylation of **36** to furnish the compound **40** was finally achieved using *N,N'*-bis(benzyloxycarbonyl)-1*H*-pyrazole-1-carboxamide in dry DMF and *N,N'*-diisopropylethylamine as base (Scheme 9a).

Scheme 8 Synthesis of 2-cyanoethyl(4-nitrophenyl)carbonate reagent **37** for -CEOC protection and its use in the preparation of guanidinylation reagents **38** and **39**



Another approach for 2'-guanidino-functionalization is to use the efficient 2'-*O*-functionalization strategy which was demonstrated in our laboratory (Chapter 3, Section A). We performed the BF₃·OEt₂-mediated electrophilic ring opening of 2,2'-anhydrouridine **3** using *N,N'*-bis(carboxybenzyl)-guanidine and *N,N'*-bis(*t*-butoxycarbonyl)-guanidine (Scheme 9, conditions vi and vii). These attempts also failed, probably due to the weaker nucleophilicity of the protected guanidines. To increase the nucleophilicity, we used the free guanidine rather than the protected guanidine and found that the reaction proceeded very smoothly. The LC-MS analysis of the crude reaction mixture confirmed the presence of the product **41**. As **41** is a highly polar compound, isolation proved challenging. A one-pot synthesis comprising ring opening of **3** with guanidine, followed by the protection of the 5'-OH group as its -DMT derivative or the protection of guanidine using base-labile groups is under investigation.

Scheme 9 Synthesis of 2'-deoxy-2'-guanidinouridine/derivative



2C.5 Significant conformational change- ¹H NMR analysis

Generally the sugar moiety in natural and modified nucleosides exists in equilibrium between *N*- and *S*-conformations. The conformation of nucleoside analogues is easily affected by electronegative substitutions in the furanose ring. To investigate whether the change of 2'-substitution from 'amino' to 'guanidino' has any effect on the sugar conformations, compound **40** was subjected to -Cbz deprotection under hydrogenation conditions to furnish free guanidine derivative **42** (Table 4).

The ¹H NMR spectra of compounds **36** and **42** were recorded and %S conformation of the furanose ring in each was calculated by a simplified equation (1),²⁷

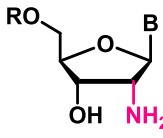
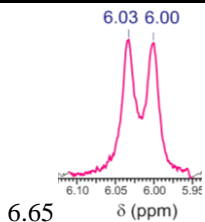
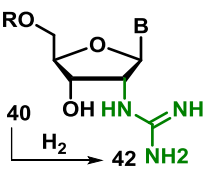
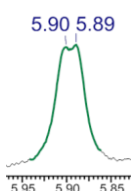
$$\%S = 100 \times (J_{H1'-H2'} - 1)/6.9 \quad (1)$$

The ¹H NMR analysis revealed a significant conformational change when 2'-amino group ($J_{H1'-H2'} = 6.65$ Hz, 82 % S conformation) was converted to the 2'-Gn group ($J_{H1'-H2'} = 2.25$ Hz, 82 % N conformation), Table 4. As anticipated, the guanidine group dictated the

Chapter 2

major N-type (C3'-endo) conformation of the sugar upon 2'-substitution at the monomeric level, which is the preferred conformation for efficient targeting of RNA.

Table 4 %S & %N calculations based on $J_{H1'-H2'}$ of 2'-amine (**36**) and 2'-guanidine (**42**) derivatives

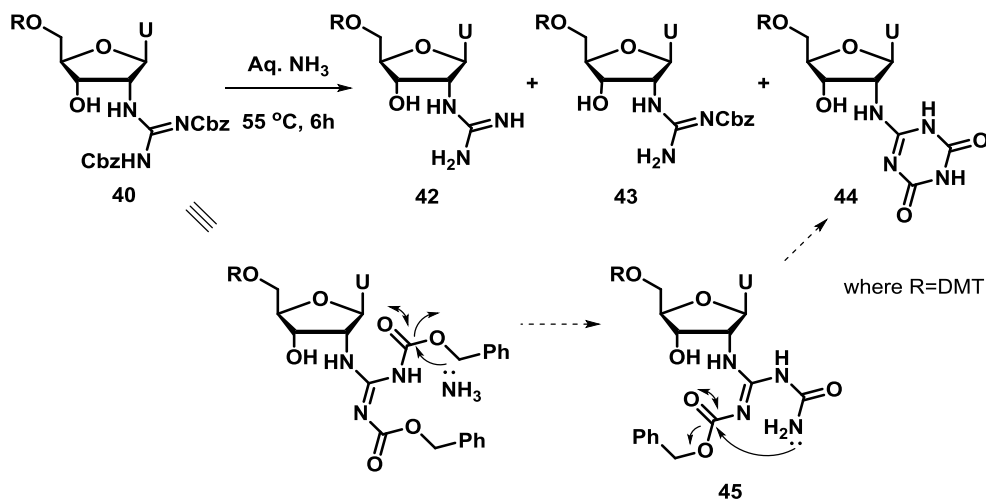
Compound (R= DMT & B= uracil)	$J_{H1'-H2'}$ (Hz) and splitting pattern	%S (C2'-endo) $100 \times (J_{H1'-H2'} - 1)/6.9$	%N (C3'-endo) (100-%S)
 36	 6.65 δ (ppm)	82	18
 40 $\xrightarrow[10\% \text{ Pd/C}]{\text{H}_2}$ 42 MeOH	 2.25 δ (ppm)	18	82

2C.6 N-Cbz deprotection under basic conditions and Synthesis of phosphoramidite

During our attempts to change the di-Cbz-protection of guanidine to base-labile protection compatible with solid phase synthesis of oligonucleotides, we realized that the –Cbz group connected to guanidine were labile to basic conditions. Before proceeding to oligonucleotides synthesis, it was necessary to check the lability of –Cbz group to deprotection conditions after solid phase oligonucleotides synthesis. In this direction we examined the stability of di-Cbz-guanidine derivative **40** to the final cleavage conditions of oligonucleotides from CPG-resin (aq. NH_3 , 55 °C, 6 h). The yield of this reaction was poor because of the poor solubility of compound **40** in aq. NH_3 . Surprisingly however, HRMS analysis of the reaction mixture after aqueous ammonia treatment of monomer **40** for 6 hours at 55 °C, showed the exclusive formation of the mono-Cbz deprotected guanidine **43**

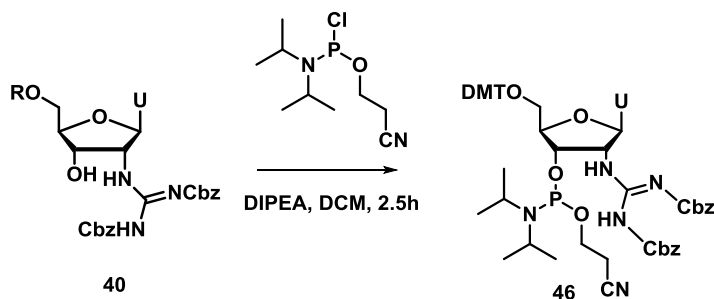
along with traces of free guanidine **42** and 2'-N-linked triazine derivative **44**. The formation of triazine derivative **44** from **40** could possibly occur through initial formation of the urea-derivative resulting from the NH_3 -mediated hydrolysis of the di-Cbz guanidine (Scheme 10), which in turn hydrolyses the second -Cbz group present on the same tether to form the 6-membered triazine in **44**.

Scheme 10 Stability of **40** to final ammonia-deprotection conditions



The products **42**, **43** and **44** obtained during this reaction are of wide interest. The N-Cbz product **43** may stabilize the duplex like any other 2'-*O*-aryl modifications⁷⁰ through additional π -stacking and may also enhance the nuclease resistance through steric blocking of the enzyme. Moreover, it could also be a good substrate (like the 2'-*O*-Bn modification carrying universal base inosine) as a high-throughput platform for RNAi.⁷¹ The triazine derivative **44** which contains hydrogen bond donors and acceptors connected to 2'-carbon through a nitrogen atom can act as a double headed nucleoside unit, similar to the report by Madsen *et al.*⁷² and with an additional methylene linker between 2'-*O*- and the nucleobase proved to destabilize the modified double headed nucleoside duplex.^{72a} We therefore, proceeded to use -Cbz protecting groups on 2'-Gn to prepare the phosphoramidite **46** (Scheme 11).

Scheme 11 Synthesis of phosphoramidite derivative **46**



2C.7 Synthesis of modified oligonucleotides, characterization and UV- T_m studies

The amidite **46** was incorporated by solid phase synthesis into the DNA oligonucleotide sequence, 5'-CCTCTTACCTCAGTTACA-3', (studied for splice correction in the previous section) as single and double modification containing oligonucleotides (Table 5, **ON23-ON26**). The modified oligonucleotides were cleaved from the resin under aq. ammonia conditions and purified by HPLC. Surprisingly, MALDI-ToF analysis of the products obtained by HPLC purification, confirmed the formation of modified oligonucleotides **ON23/ON25** and **ON24-ON26** containing 2'-*N*-Cbz-Gn (**X**, corresponding to **43**) and, 2'-*N*-triazine (**Y**, corresponding to **44**) respectively, with no traces of 2'-Gn-containing oligonucleotides. In case of double modification-containing oligonucleotide **ON25** the positions of **X** and **Y** is uncertain. Though we did not observe our designed 2'-Gn modified oligonucleotides, we were curious to see the effect of 2'-*N*-triazine (a potential double-headed nucleoside) on the target recognition. Hence, we undertook the biophysical studies of the modified oligonucleotides obtained.

Table 5 2'-modified oligonucleotides synthesized, MALDI-ToF characterized and their UV- T_m data.

Code	Sequences (5'-3') [†]	Mass	UV T_m (ΔT_m) in °C*
------	--------------------------------	------	----------------------------------

		Calcd/Obsd	cDNA	cRNA
ON1	CCTCTTACCTCAGTTACA	5366.9/5367.2	54.7	56.6
ON23	CCTCTTACCTCAGT <u>X</u> ACA	5543.9/5543.0	51.1(-3.6)	54.1(-2.5)
ON24	CCTCTTACCTCAGT <u>Y</u> ACA	5478.9/5474.6	50.3(-4.4)	53.0(-3.6)
ON25	CCTCTTACC <u>X</u> CAGT <u>Y</u> ACA	5655.9/5654.0	45.6(-9.1)	49.4(-7.2)
ON26	CCTCTTACC <u>Y</u> CAGT <u>Y</u> ACA	5590.9/5585.5	44.4(-10.3)	48.4(-8.2)
ON17	ccucuuaccucagu <u>a</u> ca	5823.0/5824.4	53.4	77.0
ON27	ccucuuaccucagu <u>X</u> aca	5984.0/5980.4	49.7(-3.7)	72.7(-4.3)
ON28	ccucuuaccucagu <u>Y</u> aca	5919.0/5932.0	45.2(-8.2)	70.7(-6.3)

[†]X=2'-(*N*-Cbz)-Gn and Y=2'-*N*-triazine; *UV- T_m values were measured by annealing 1 μ M each sequence with 1 μ M cDNA in sodium phosphate buffer (0.01M, pH 7.2) containing 100mM NaCl and are averages of three independent experiments. Lower case represents 2'-OMe nucleotides

The UV- T_m studies of the modified oligonucleotides (**ON23-ON26**) containing both 2'-*N*-Cbz-Gn (**X**) and 2'-*N*-triazine (**Y**) substituted oligonucleotides were found to destabilize the duplexes with cDNA and cRNA (Table 5 and Figure 18). However, the destabilization is relatively less in oligonucleotides **ON23** and **ON25** than in oligonucleotides **ON24** and **ON26**, respectively. The observed difference in destabilization of **X** and **Y**-containing modified oligonucleotides could easily be attributed to the presence of the 'Cbz' protecting group in **ON23** and **ON25**, due to the possible π - π stacking strength as in 2'-*O*-aryl modified AONs.⁷⁰

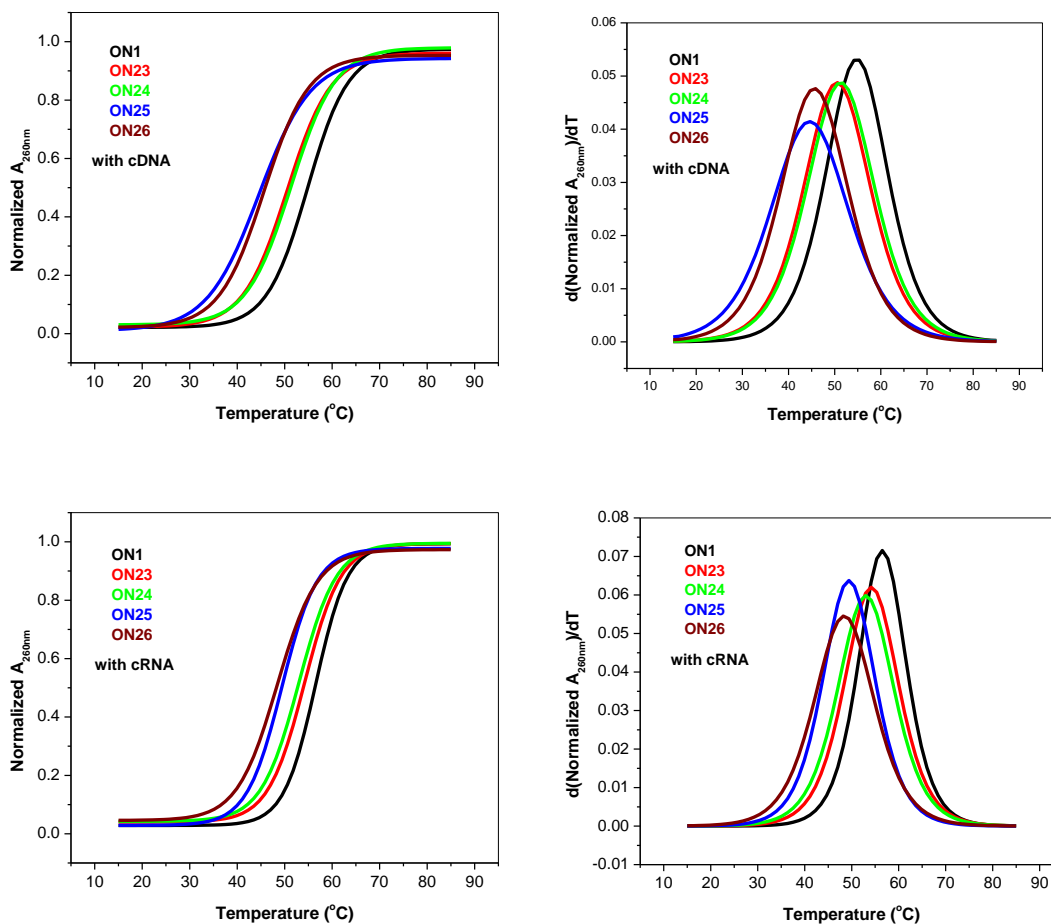


Figure 18 UV melting profiles and first derivative curves of complexes of modified oligonucleotides with cDNA or cRNA

These results with modified DNA oligonucleotides prompted us to examine whether these modifications would have a similar effect when incorporated in 2'-OMe RNAs. Accordingly, we synthesized 2'-OMe RNAs incorporating single unit of amidite **46**. Similar to the modified DNA oligonucleotides, we obtained a ~1:1 ratio of **X** and **Y**-containing modified oligonucleotides **ON27** and **ON28**, respectively (Please see experimental section for crude HPLC chromatogram). The UV- T_m studies of **ON27** and **ON28** again revealed the destabilization of the duplexes with both cDNA and cRNA. The degree of destabilization was more when these modifications were incorporated in 2'-OMe RNA than in DNA.

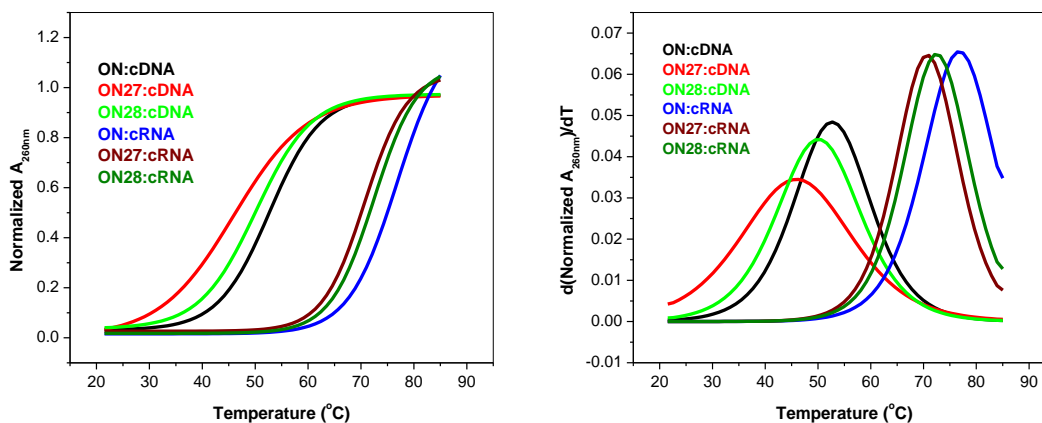


Figure 18 UV melting profiles and first derivative curves of complexes of modified oligonucleotides with cDNA or cRNA

2C.8 Conclusion

- The guanidinylation of 2'-deoxy-2'-aminouridine proved to be challenging and was achieved with *N,N'*-bis(benzyloxycarbonyl)-1*H*-pyrazole-1-carboxamide in moderate yield.
- Our design of 2'-Gn modification to play a synergistic role resulted in the serendipitous formation of the 2'-*N*-triazine derivative which has potential as a double-headed nucleoside.
- The 2'-*N*-triazine and 2'-(*N*-Cbz)-Gn modification-containing oligonucleotides destabilize the duplexes with cDNA and cRNA, the degree of destabilization being less in (*N*-Cbz)-Gn oligonucleotides, probably due to the stabilizing stacking interactions offered by the -Cbz group.

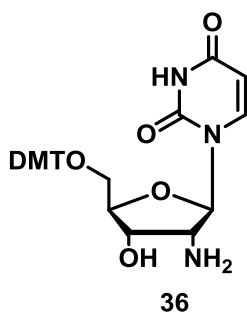
2C.9 Experimental Section

General

All the non-aqueous reactions were carried out under the inert atmosphere of Nitrogen/Argon and the chemicals used were of laboratory or analytical grade. All solvents used were dried and distilled according to standard protocols. TLCs were carried out on pre-coated silica gel GF254 sheets (Merck 5554). All reactions were monitored by TLC and usual work-up implies sequential washing of the organic extract with water and brine followed by drying over anhydrous sodium sulfate and evaporation of the organic layer under vacuum. Column chromatographic separations were performed using silica gel 100-200 mesh (Merck) or 230-400 mesh (Merck) and using the solvent systems EtOAc/Petroleum ether or MeOH/DCM. ^1H and ^{13}C NMR spectra were obtained using Bruker AC-200, AC-400 or AC-500 NMR spectrometers. The chemical shifts are reported in delta (δ) values and referred to internal standard TMS for ^1H . LC-MS and HRMS-mass spectra were recorded on a Finnigan-Matt mass spectrometer. Complementary DNA sequences were synthesized on Bioautomation Mer-Made 4 DNA synthesizer using standard β -cyanoethyl phosphoramidite chemistry. DNA sequences were also analyzed and purified under the same conditions with the increasing gradient of acetonitrile in 0.1N triethylammonium acetate of pH 7.0. The MALDI-TOF spectra were recorded on AB Sciex TOF/TOFTM series explorerTM 72085 instrument; THAP (2, 4, 6-trihydroxyacetophenone) was used as the matrix for DNA samples characterization. Thermal denaturation experiments were performed on Varian Cary-300 UV-Vis spectrophotometer fitted with a peltier-controlled temperature programmer and a water circulator, at the temperature ramping rate 0.5 $^\circ\text{C}/\text{min}$ and the absorbance was recorded at 260 nm for every 0.5 $^\circ\text{C}$ rise in temperature.

5'-O-DMT-2'-amino-2'-deoxyuridine, 36:

Synthesis of **36** was achieved using the one pot synthesis protocol from uridine reported by McGee *et al.*⁶⁷ To anhydrouridine **3** (22.12 mmol, 5.0 g), 4,4'-dimethoxytrityl chloride (23.23 mmol, 7.86 g) and DMAP (22 mg, catalytic), was added a mixture of dry pyridine



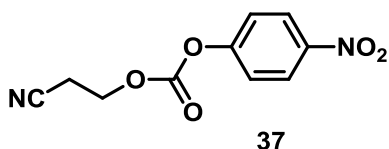
(18 mL) and dry DMF (14 mL), and the reaction was stirred for 16 h at rt. Solvents were removed partially under reduced pressure. The concentrated reaction mixture was partitioned between dichloromethane and water. 10% aqueous NaHCO₃, water and brine solution washes were given to the organic layer. The organic layer was dried over anhydrous Na₂SO₄ and concentrated to dryness. The residue was co-evaporated twice with toluene to afford a sticky oil.

To this reddish oil were added trichloroacetonitrile (25 mL) and triethylamine (0.7 mL), and the reaction mixture was refluxed for 16 h. The black reaction mixture was then concentrated using rota evaporator and LC-MS analysis of resulting black oil confirmed the formation of oxazoline **35**. To a solution of oxazoline **35** in ethanol (60 mL) was added a 6 N sodium hydroxide solution (30 mL), and the reaction mixture was refluxed for 16 h, then cooled, and evaporated partially. The residue was partitioned between dichloromethane and saturated ammonium chloride solution. The aqueous phase was again washed with dichloromethane, and the combined organic layers were dried over anhydrous Na₂SO₄, filtered, and evaporated to dryness to give yellowish foam. The foam was purified by chromatography on silica gel to furnish compound **36** (eluted in 5-10% MeOH in DCM containing 1% NEt₃) in 60% overall yield as pale yellow foam. **35** LC-MS (EI): Mass calculated for C₃₂H₂₈Cl₃N₃O₇Na (M+Na) 695.92, found 696.23. **36** ¹H NMR (DMSO-d₆, 200 MHz): δ 3.02-3.13 (m, 2H), 3.75 (s, 6H), 3.89 (t, 1H, J= 6.76Hz), 4.07-4.11 (m, 1H), 4.33 (bs, 1H), 4.57 (t, 1H, J= 5.2 Hz), 5.54 (d, 1H, J= 7.7 Hz), 6.01 (d, 1H, J= 6.6 Hz), 6.31 (bs, 1H), 6.89-6.93 (m, 4H), 7.23-7.37 (m, 9H), 7.59 (d, 1H, J= 7.82 Hz); ¹H NMR (DMSO-d₆, D₂O, 200 MHz): δ 3.03-3.14 (m, 2H), 3.74 (s, 6H), 3.87 (t, 1H, J= 6.5Hz), 4.09 (m, 1H), 4.31 (m, 1H), 5.54 (d, 1H, J= 8.09 Hz), 5.98 (d, 1H, J= 6.19 Hz), 6.88-6.92

(m, 4H), 7.22-7.35 (m, 9H), 7.58 (d, 1H, J= 8.09 Hz); **LC-MS (EI)**: Mass calculated for C₃₀H₃₁N₃O₇Na (M+Na) 568.20, found 568.23.

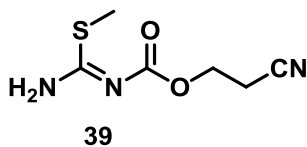
2-cyanoethyl(4-nitrophenyl)carbonate, 37:

To a solution of 2-cyanoethanol (70.3 mmol, 5 g) in dry DCM and dry pyridine (105.4 mmol, 8.33 mL) was added *p*-nitrophenylchloroformate (77.3 mmol, 16 g) at 0 °C and then the reaction mixture was stirred at rt for 4 h. The reaction mixture was further diluted with DCM and partitioned by the addition of minimum volume of water. The aqueous layer was again extracted with DCM and the combined organic layers were washed with 5% aq. NaHCO₃ solution, brine solution and dried over anhydrous Na₂SO₄. DCM was removed on rota evaporator and the crude compound was column purified (eluted in 20% EtOAc in petroleum ether) to yield reagent **37** as a white solid in 88% yield. **¹H NMR** (CDCl₃, 200 MHz): δ 2.86 (t, 2H, J= 6.28 Hz), 4.50 (t, 2H, J= 6.28 Hz), 7.38 (d, 2H, J= 9.12 Hz), 8.28 (d, 2H, J= 9.12 Hz).



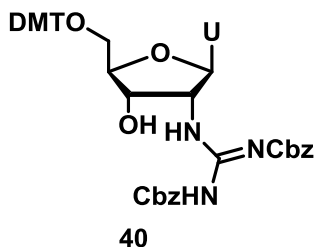
Mono-(2-cyanoethoxycarbonyl)-protected guanidylating agents, 38/39:

To a mixture of *S*-methylthiourea sulphate/*1H*-pyrazole-1-carboxamide hydrochloride and sodiumbicarbonate in DCM was added activated carbonate of 2-cyanoethanol **37**. The reaction mixture was stirred at rt overnight. The reaction mixture was further diluted with DCM and partitioned with water. To the organic layer brine wash was given. The organic layer was further dried over anhydrous Na₂SO₄ and concentrated under reduced pressure. The crude compound was column purified to obtain the reagents as white solids. **38** **¹H NMR** (CDCl₃, 200 MHz): δ 2.79 (t, 2H, J= 6.74 Hz), 4.38 (t, 2H, J= 6.74 Hz), 6.43 (dd, 1H, J= 1.1 and 2.72 Hz), 6.89 (d, 1H, J= 9.1 Hz), 7.71 (d, 1H, J= 1.13 Hz), 8.10 (d, 1H, J= 9.17 Hz), 8.40 (d, 1H, J= 2.72 Hz). **39** **¹H NMR** (CDCl₃, 200 MHz): δ 2.47 (s, 3H), 2.76 (t, 2H, J= 6.74 Hz), 4.32 (t, 2H, J= 6.74 Hz).



5'-O-DMT-2'-deoxy-2'-[N,N'-bis(benzyloxycarbonyl)guanidino]uridine, 40:

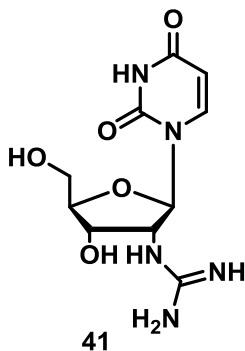
Compound **36** (2.2 mmol, 1.2 g) was dissolved in dry DMF (5 mL) and DIPEA (6.6 mmol, 1.14 mL) was added. To the reaction mixture *N,N'*-bis(benzyloxycarbonyl)-1*H*-pyrazole-1-carboxamide (6.6 mmol, 2.5 g) was added and the reaction mixture was stirred at rt for 4 h. Solvents were partially removed under reduced pressure and the reaction mixture was diluted with EtOAc. The organic



layer was washed with water, followed by brine solution. The organic layer was separated and dried over anhydrous Na₂SO₄. The crude compound was column purified to provide compound **40** (eluted in 55% EtOAc in petroleum ether) as white foam in 71% yield. ¹H NMR (CDCl₃, 200 MHz): δ 3.37-3.38(m, 2H), 3.7 (s, 6H), 4.13 (m, 1H), 4.49 (m, 1H), 4.78-4.88 (m, 1H), 5.03(s, 2H), 5.10 (s, 2H), 5.39 (d, 1H, J= 8.08 Hz), 6.13 (d, 1H, J= 7.45 Hz), 6.77 (m, 4H), 7.17-7.4 (m, 9H), 7.69 (d, 1H, J= 8.1 Hz), 11.58 (s, 1H); ¹H NMR (CDCl₃, D₂O, 200 MHz): δ 3.38(m, 2H), 3.7 (s, 6H), 4.15 (m, 1H), 4.45-4.49 (m, 1H), 4.79 (m, 1H), 5.03(s, 2H), 5.10 (s, 2H), 5.39 (d, 1H, J= 8.07 Hz), 6.12 (d, 1H, J= 7.32 Hz), 6.77 (m, 4H), 7.21-7.40 (m, 9H), 7.69 (d, 1H, J= 8.2 Hz); ¹³C NMR (CDCl₃, 50 MHz): δ 55.0, 58.8, 63.6, 66.9, 68.2, 71.1, 85.0, 86.8, 87.2, 102.4, 113.1, 126.8, 127.8, 128.3, 128.4, 129.9, 130.0, 134.2, 135.0, 136.3, 140.9, 144.3, 150.9, 153.1, 155.7, 158.4, 163.3, 163.4; HRMS (EI): Mass calculated for C₄₇H₄₅O₁₁N₅Na (M+Na) 878.3008, found 878.3036.

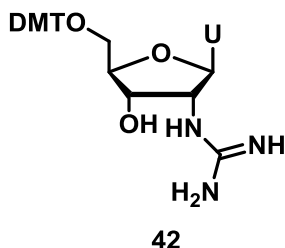
2'-deoxy-2'-guanidinouridine, **41**:

Desiccated 2,2'-anhydrouridine **3** (100 mg, 0.44 mmol,) was dissolved in dry DMA (3 mL) followed by the addition of BF₃.OEt₂ (1.1 mmol, 136 μL) under argon atmosphere. After 2 minutes, the neutralized guanidine hydrochloride (with 0.45 M NaOEt) in ethanol (3 mL) was added and the resulting mixture was stirred at 130 °C for 8-12 h. TLC analysis indicated the consumption of starting material and the formation of a polar compound. The LCMS analysis confirmed the formation of 2'-deoxy-2'-guanidinouridine. **41** LCMS (EI): Mass calculated for C₁₀H₁₆N₅O₅ (M+H) 286.11, found 286.09.



5'-O-DMT-2'-deoxy-2'-guanidinouridine, 42:

Compound **40** was dissolved in 5mL MeOH followed by the addition of 10% Pd-C (10%

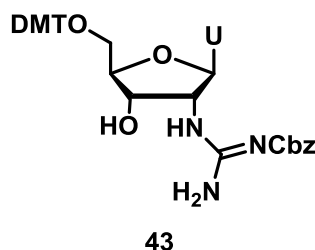


w/w, 0.18 g). The reaction mixture was subjected to catalytic hydrogenation at 50 psi of hydrogen pressure for 2 h. After the TLC analysis, reaction mixture was filtered over celite and the removal of methanol in *vacuo* gave the free amine. The product was precipitated by adding EtOAc. **Yield??** $^1\text{H NMR}$ (DMSO-d₆, 500 MHz): δ 3.25-3.28 (m, 2H), 3.74 (s, 6H), 4.00-4.10 (m, 2H),

4.27 (m, 1H), 5.26 (d, 1H, J= 7.13 Hz), 5.86 (d, 1H, J= 2.0 Hz), 6.90 (m, 4H), 7.25-7.42 (m, 9H), 7.60 (d, 1H, J= 7.1 Hz). $^1\text{H NMR}$ (DMSO-d₆, D₂O, 500 MHz): δ 3.26-3.29 (m, 2H), 3.72 (s, 6H), 3.99-4.06 (m, 2H), 4.24 (t, 1H, J = 6.1 Hz), 5.25 (d, 1H, J= 8.03 Hz), 5.84 (d, 1H, J= 2.4 Hz), 6.87 (m, 4H), 7.25-7.37 (m, 9H), 7.54 (d, 1H, J= 7.94 Hz); $^{13}\text{C NMR}$ (DMSO-d₆, 125 MHz): δ 55.1, 57.5, 62.5, 67.6, 82.3, 85.8, 87.6, 102.5, 113.3, 126.8, 127.7, 127.9, 129.82, 129.85, 135.1, 135.4, 144.7, 158.0, 158.1, 159.7; **HRMS (EI)**: Mass calculated for C₃₁H₃₄O₇N₅ (M+H) 588.2453, found 588.2455.

5'-O-DMT-2'-deoxy-2'-(N-benzyloxycarbonyl)guanidinouridine, 43:

Compound **40** (0.11 mmol, 100 mg) was taken in a Wheaton vial and aqueous ammonia (2



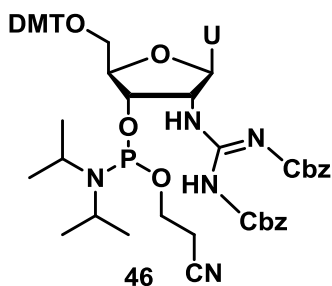
mL) was added. It was kept in a water bath at 55 °C for 6 h, similar to the cleavage conditions of oligonucleotides from the solid support. As the solubility of **40** was poor in aqueous media, the conversion was not clean. The LCMS analysis of the crude reaction mixture (spectrum included) revealed a major peak

corresponding to **43** and two minor peaks corresponding to free guanidine derivative **42** and a possible triazine derivative **44**. The identity of the compounds was further confirmed through HRMS analysis. The isolation of the individual compounds from the reaction mixture proved difficult. Pure compound **43** was obtained through column chromatography (eluted in 5% MeOH in EtOAc) and was further characterized by ^1H and ^{13}C NMR. **42** **HRMS (EI)**: Mass calculated for C₃₁H₃₄O₇N₅ (M+H) 588.2453, found 588.2455; **44**

HRMS (EI): Mass calculated for $C_{33}H_{32}O_9N_6Na$ (M+Na) 679.2123, found 679.2120; **1H NMR** (Acetone- d_6 , 500 MHz): δ 3.31 (s, 1H), 3.36-3.55 (m, 3H), 3.74 (s, 6H), 4.25 (m, 1H), 4.65 (s, 1H), 5.06 (s, 1H), 5.43 (s, 1H), 6.17 (s, 1H), 6.94 (m, 4H), 7.20-7.38 (m, 9H), 7.85 (s, 1H); **^{13}C NMR** (Acetone- d_6 , 125 MHz): δ 49.8, 55.6, 57.8, 62.5, 64.6, 71.8, 86.3, 87.7, 103.2, 114.2, 127.7, 128.9, 131.0, 131.1, 136.3, 136.6, 141.5, 145.4, 152.3, 159.7, 164.2, 173.6; **HRMS (EI):** Mass calculated for $C_{39}H_{40}O_9N_5$ (M+H) 722.2821, found 722.2823.

Phosphoramidite derivative, **46**:

To the compound **40** (0.58 mmol, 500 mg) dissolved in dry DCM (10 mL), DIPEA (1.75 mmol, 0.34 mL) was added. 2-cyanoethyl-*N,N*-diisopropylchlorophosphine (1.17 mmol, 0.26 mL) was added to the reaction mixture at 0 °C and stirring continued at room temperature for 2.5 h. The contents were diluted with DCM and washed with 5% $NaHCO_3$ solution. The organic phase was dried over anhydrous Na_2SO_4 and concentrated to white

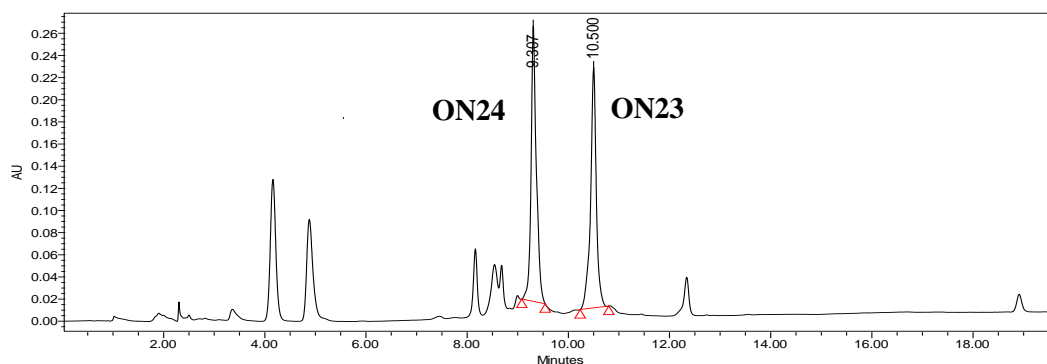


foam. The residue was re-dissolved in DCM and the compound was precipitated with petroleum ether to obtain corresponding phosphoramidite derivative **46** in 82 % yield. **^{31}P NMR** (Acetonitrile, D_2O as external standard, 400 MHz): δ 149.78, 150.24; **HRMS (EI)** Mass calculated for $C_{56}H_{62}O_{12}N_7PNa$ (M+Na) 1078.4086, found 1078.4072.

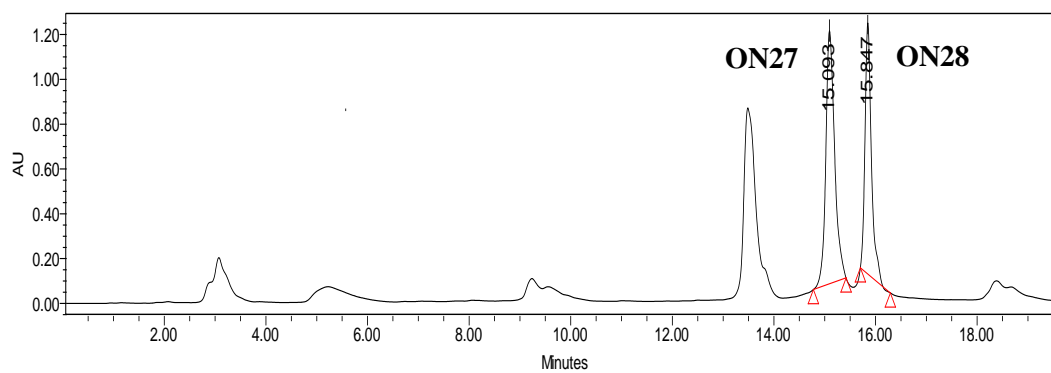
2C.10 Appendix

Compounds - Spectral data	Page No.
Representative crude HPLC profiles	154
36 - ¹ H NMR & D ₂ O exchange, 37 - ¹ H NMR	155
38, 39, 40 - ¹ H NMR	156
40 D ₂ O exchange, ¹³ C NMR & DEPT NMR	157
42 - ¹ H NMR & D ₂ O exchange & ¹³ C NMR	158
42 - DEPT, 43 - ¹ H NMR & ¹³ C NMR	159
43 - DEPT & 46 - ¹³ C NMR	160
35 & 36 - LC MS	161
40 -HRMS & 41 - LC MS	162
Crude LC MS of Compound 40 after NH ₃	163
43 & 44 - HRMS	164
46 - HRMS & ON23 -MALDI-TOF spectra	165
ON24 & ON25 -MALDI-TOF spectra	166
ON26 & ON27 -MALDI-TOF spectra	167
ON28 -MALDI-TOF spectra	168

Representative crude HPLC profiles



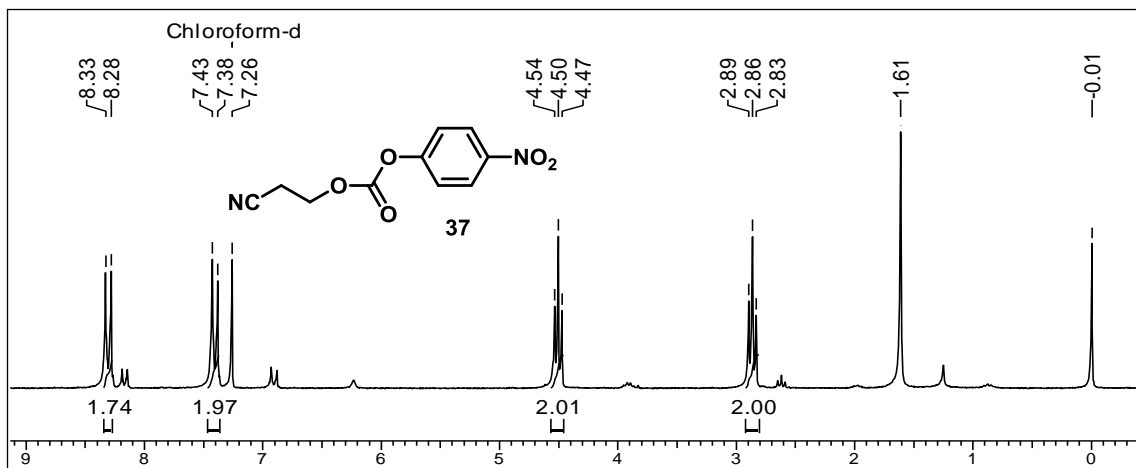
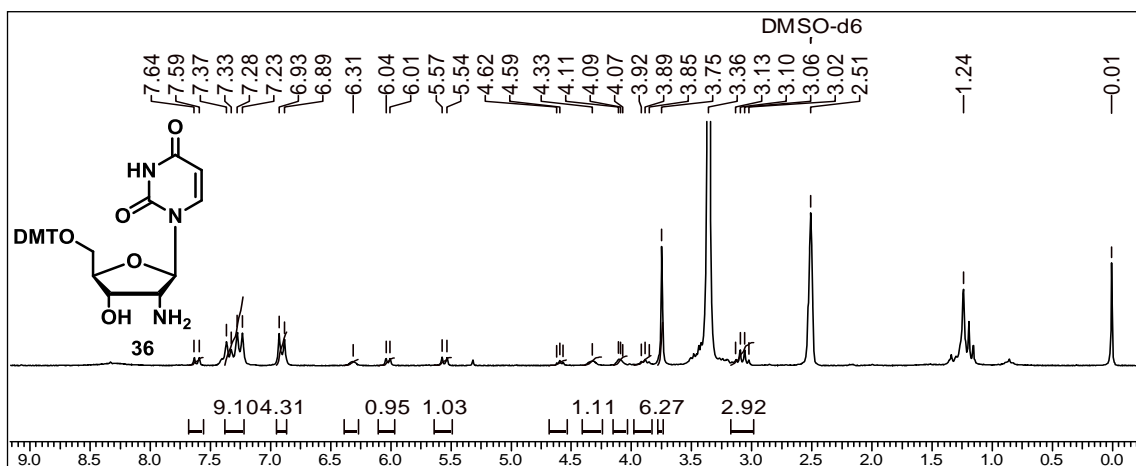
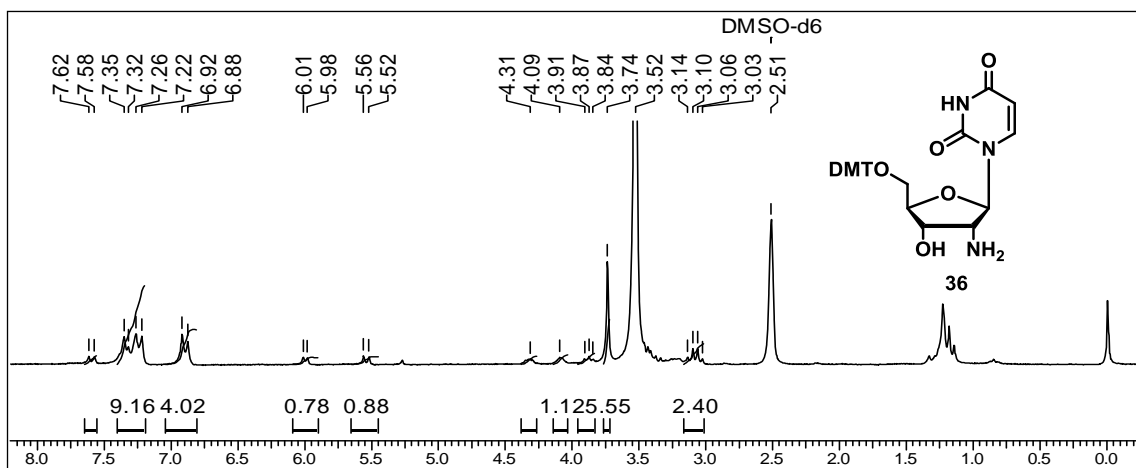
Crude HPLC profile of single unit of amidite **46** incorporated in DNA oligonucleotides

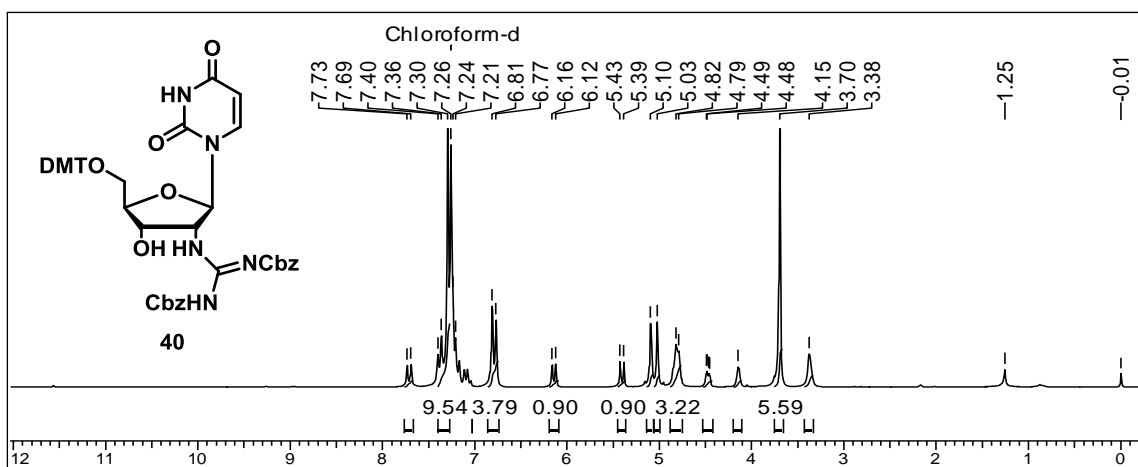
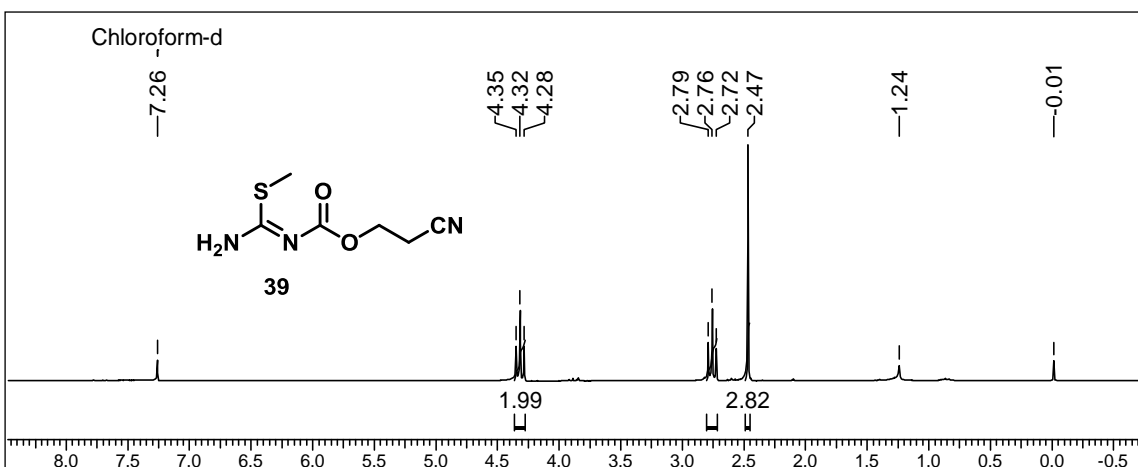
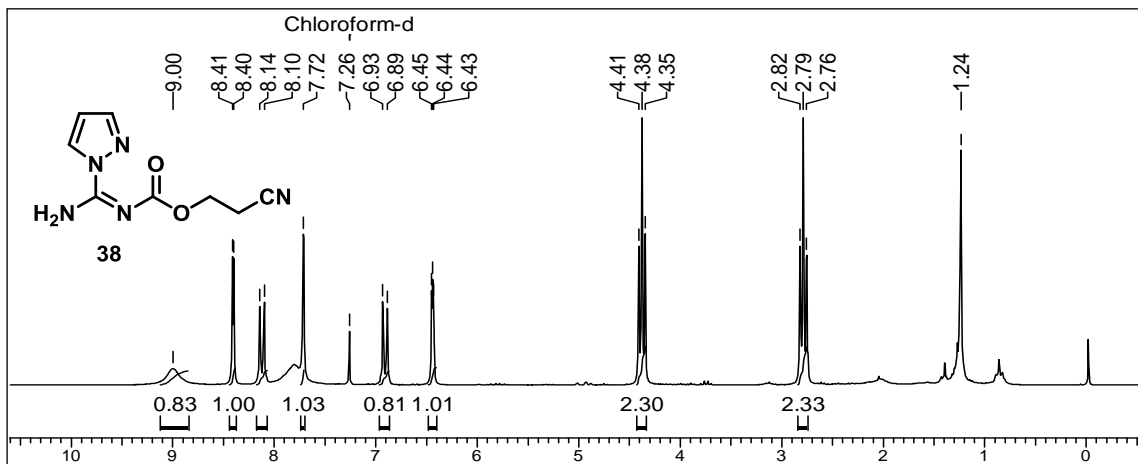


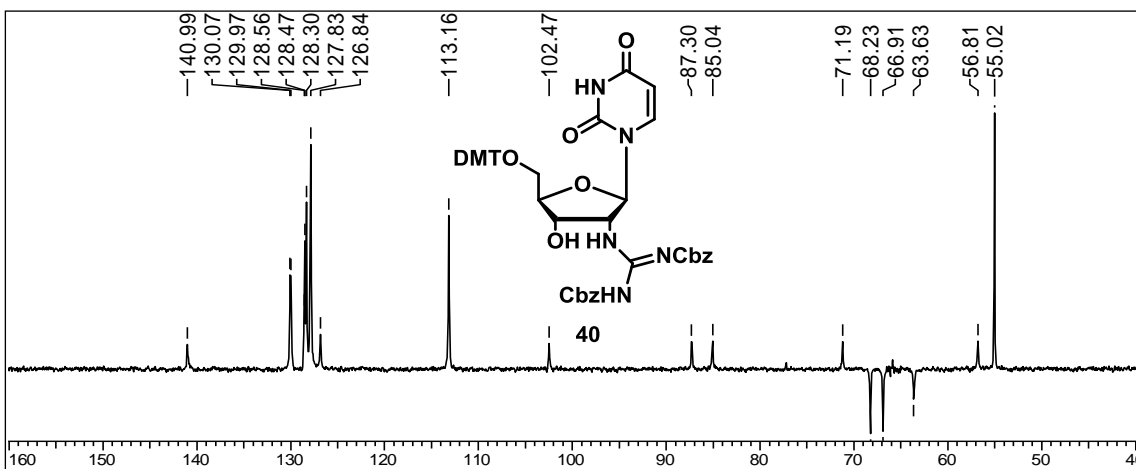
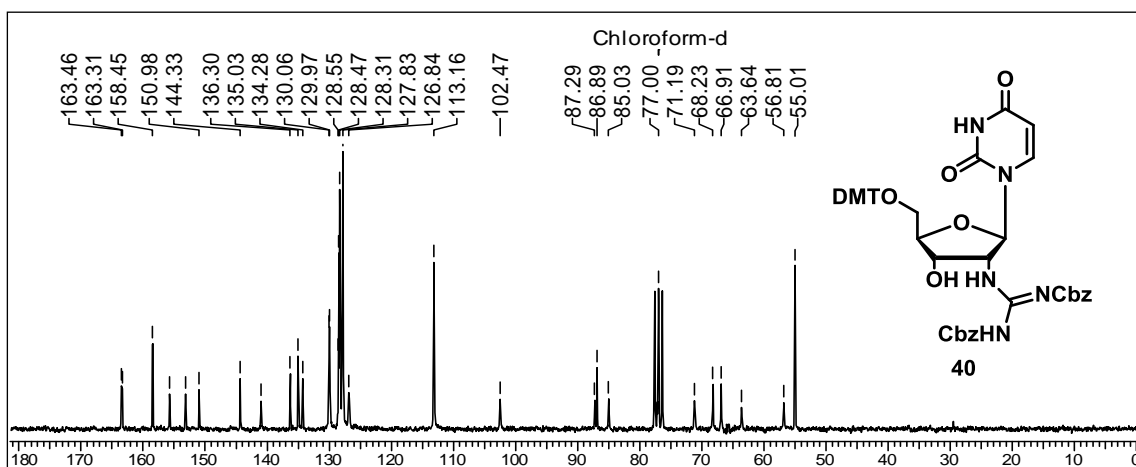
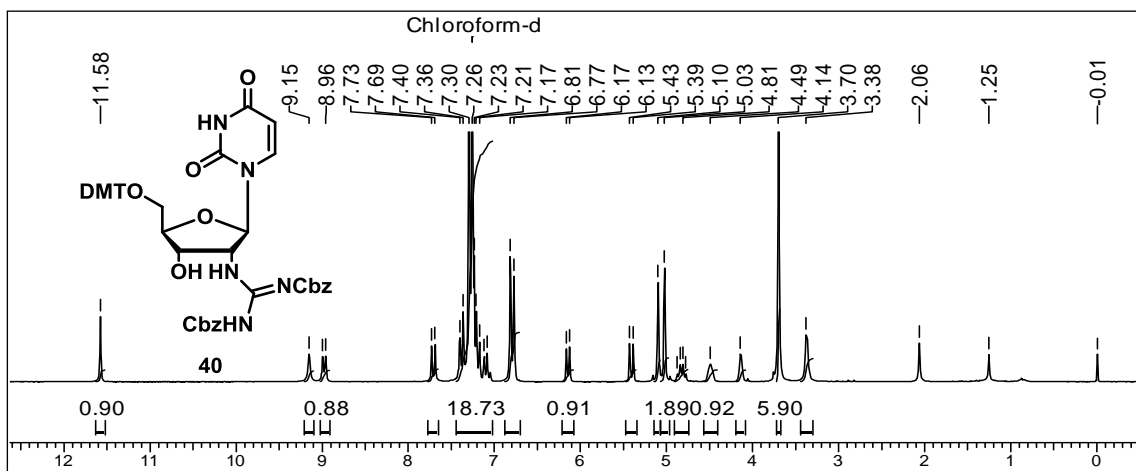
Crude HPLC profile of single unit of amidite **46** incorporated in 2'-OMe RNA oligonucleotide

Chapter 2

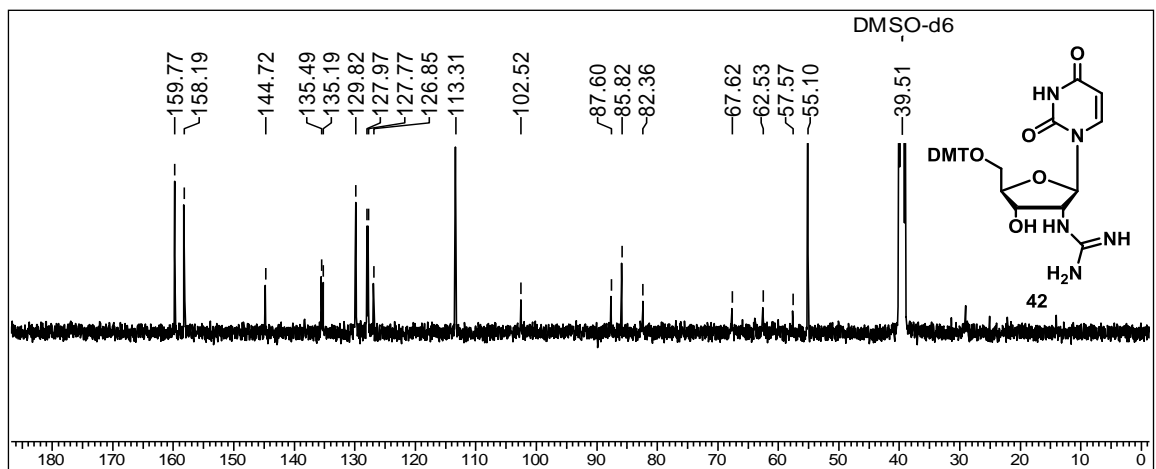
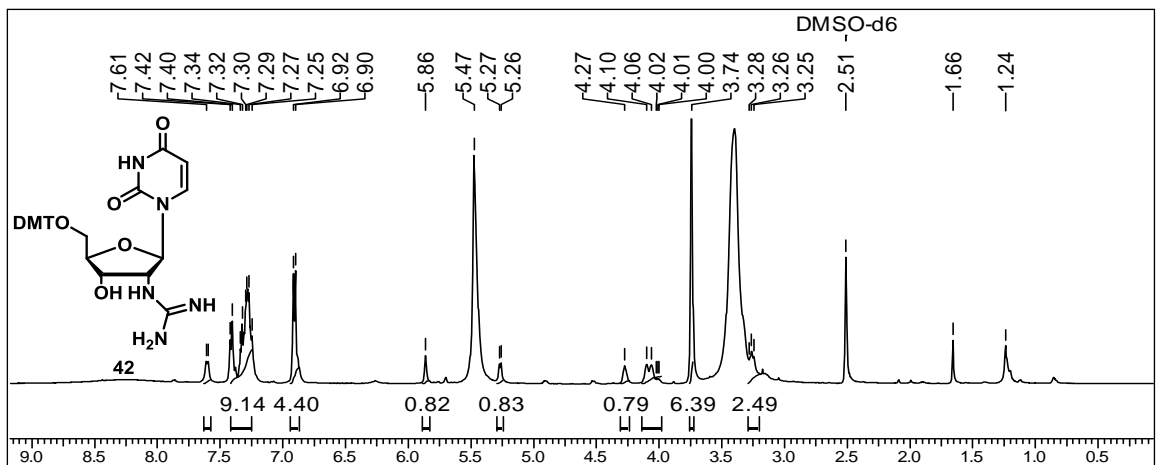
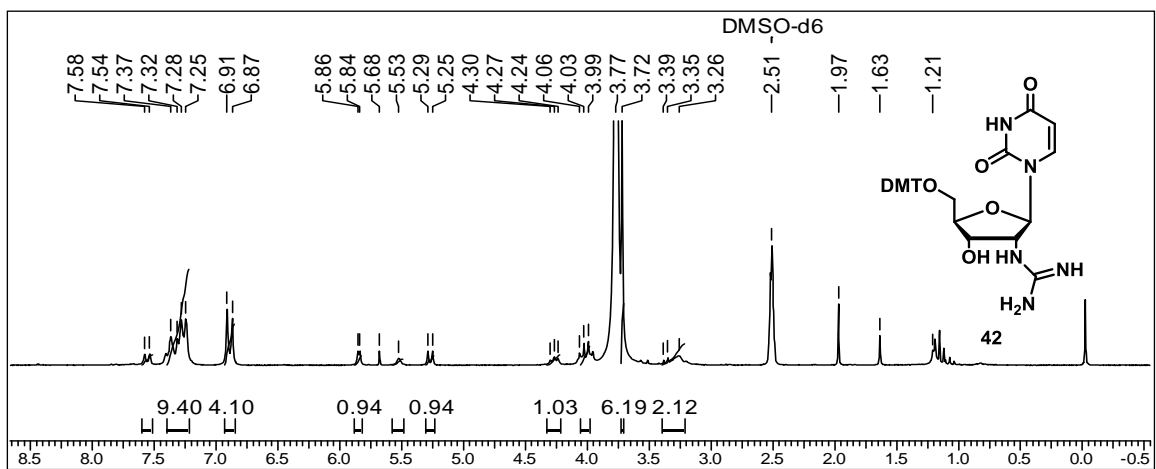
NMR spectral data:



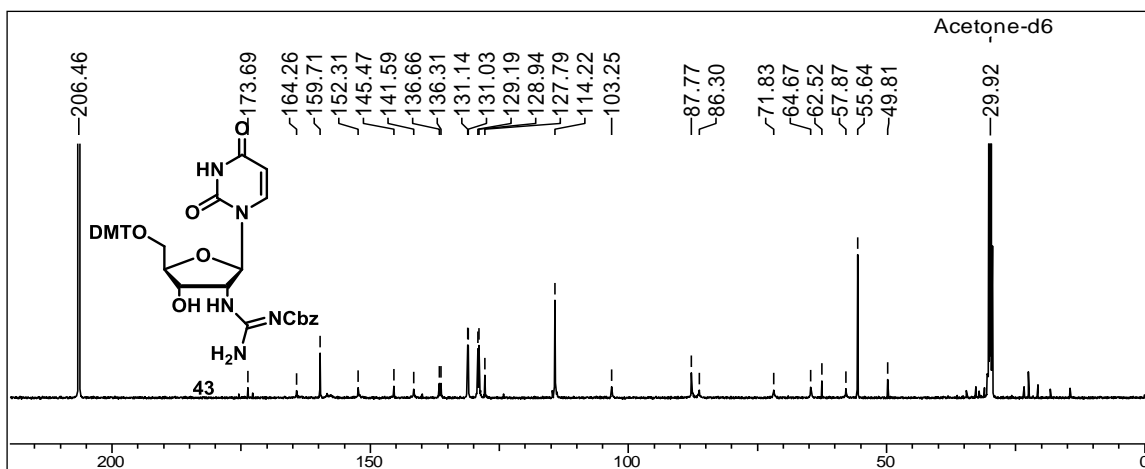
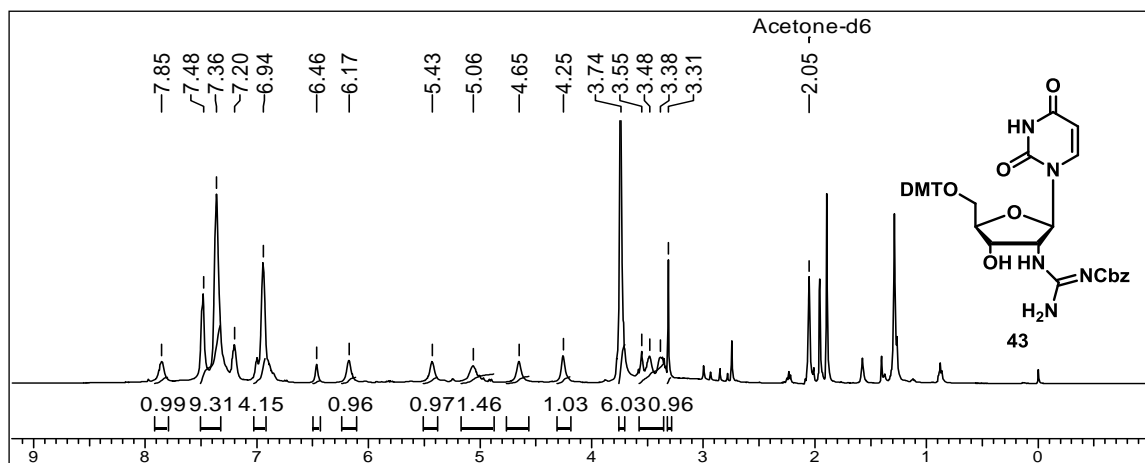
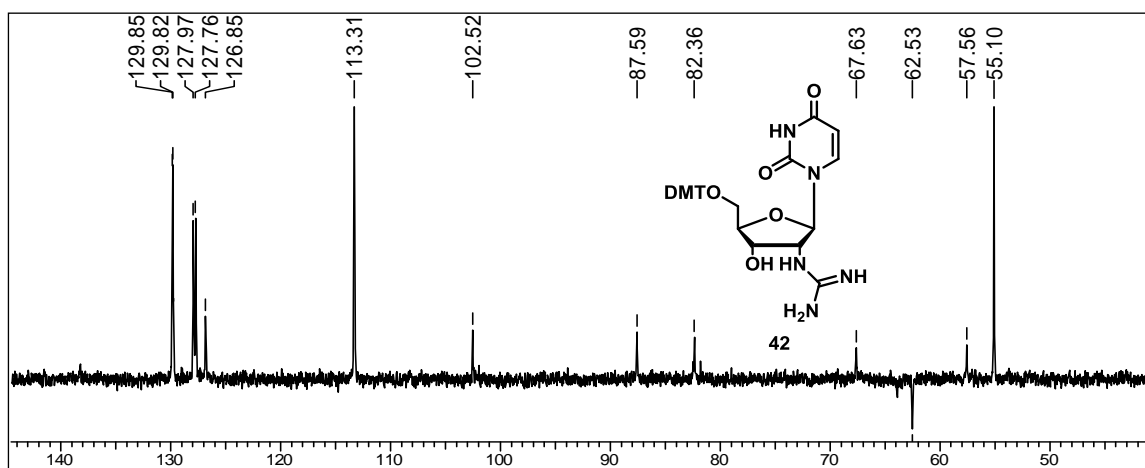




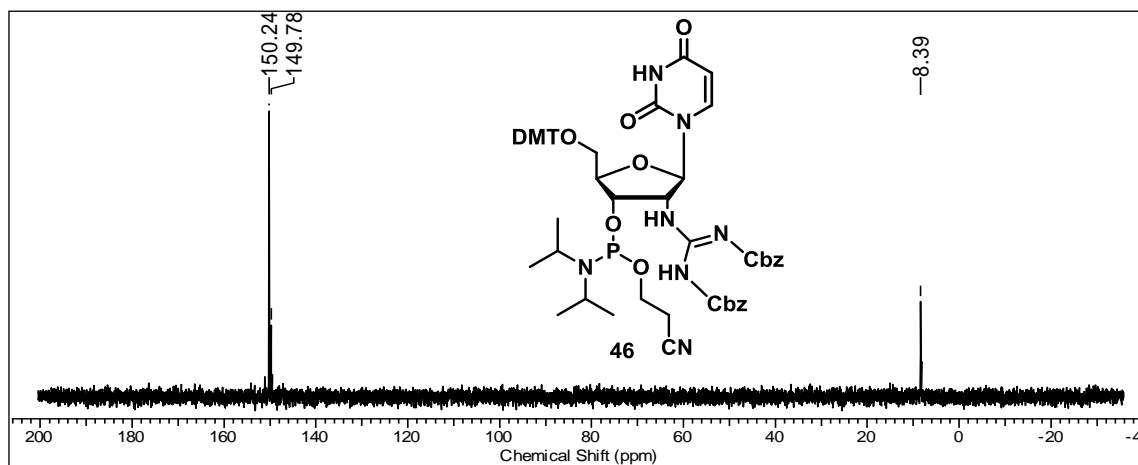
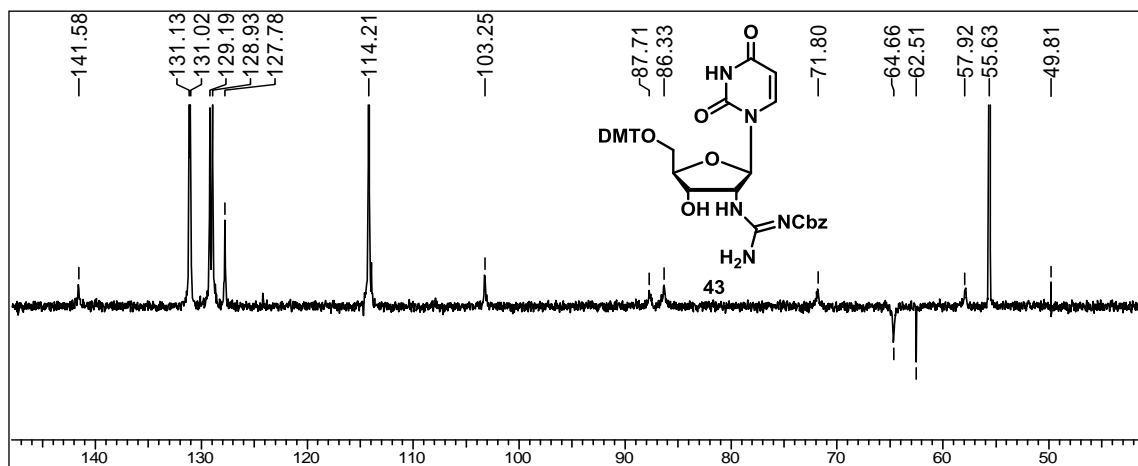
Chapter 2



Chapter 2

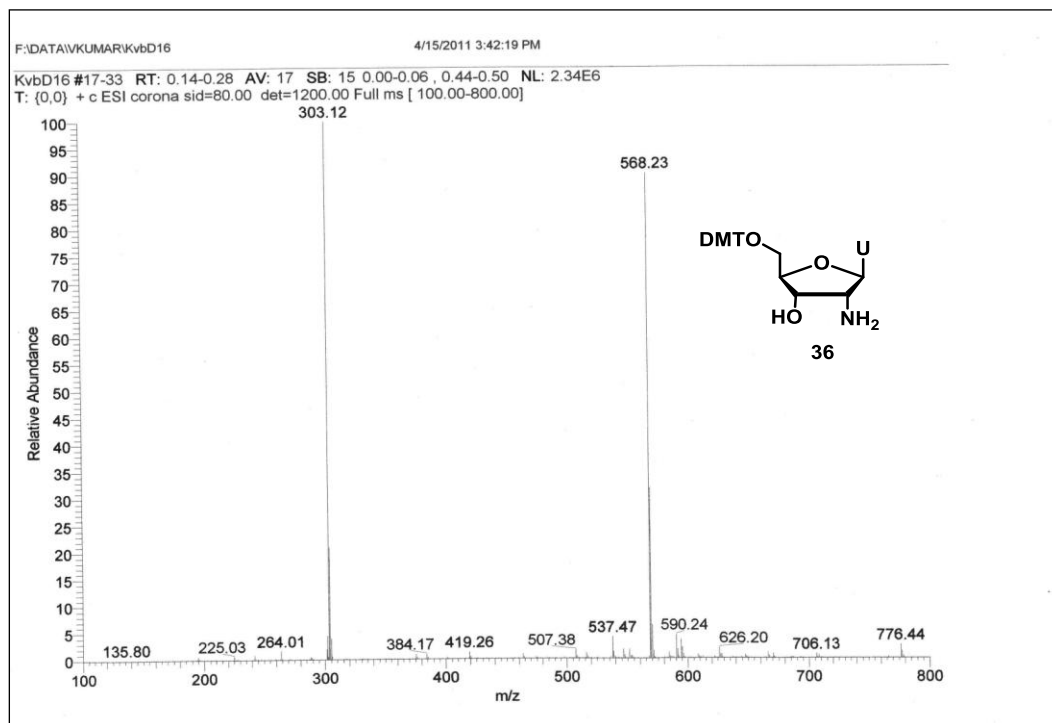
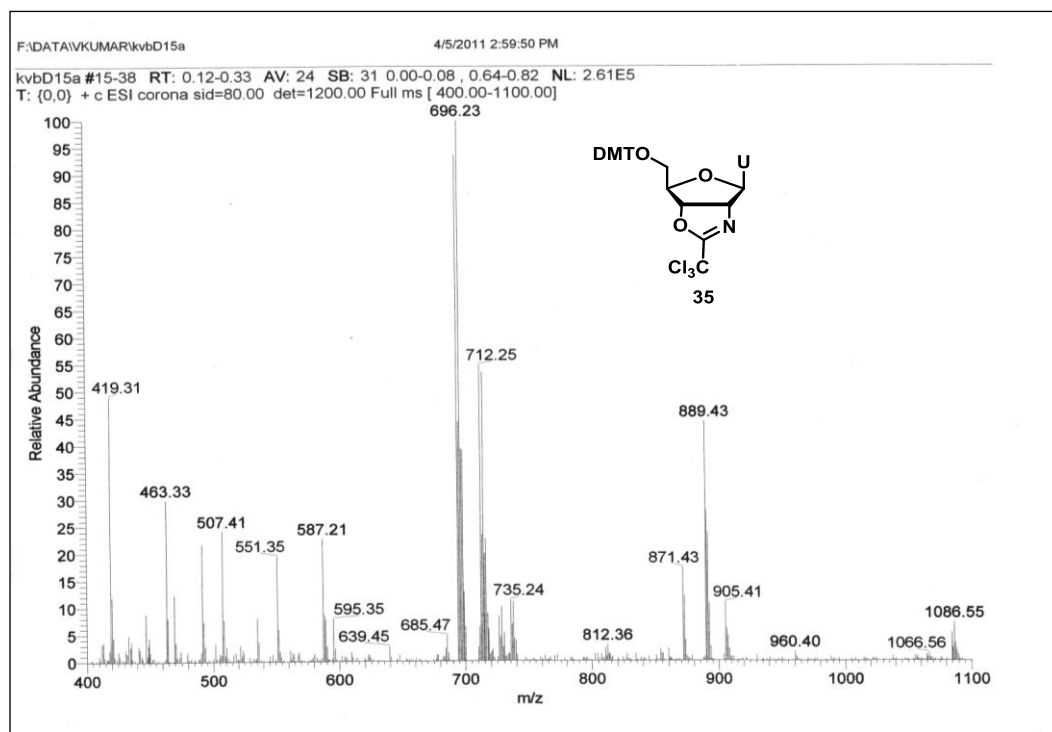


Chapter 2

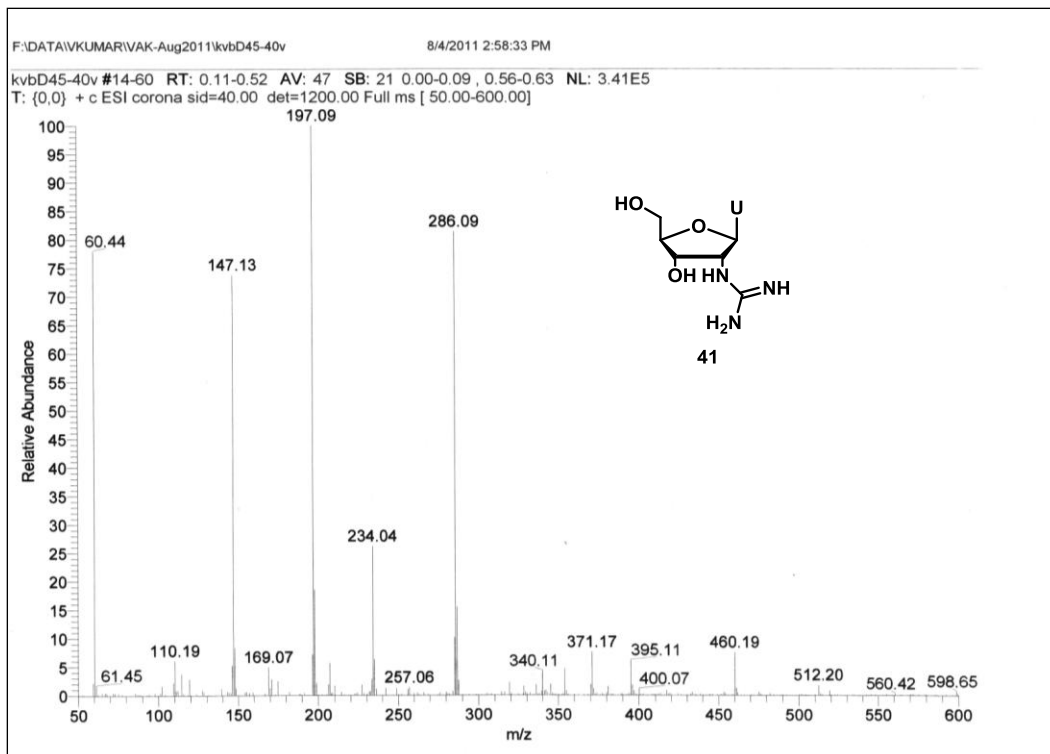
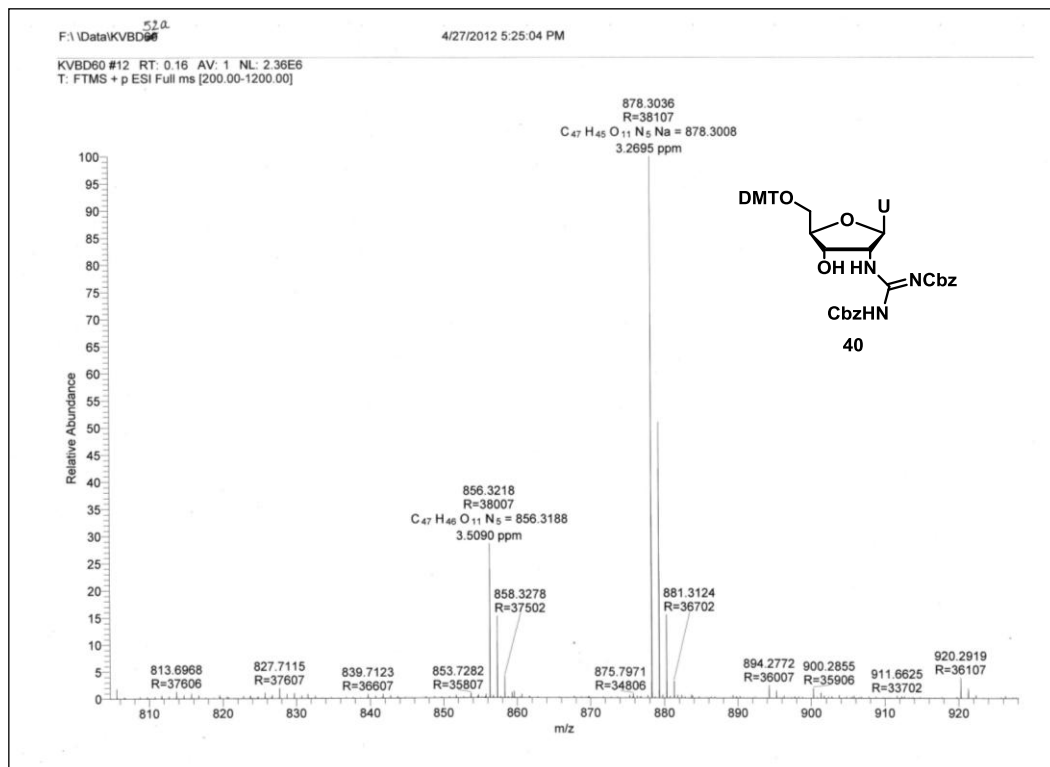


Chapter 2

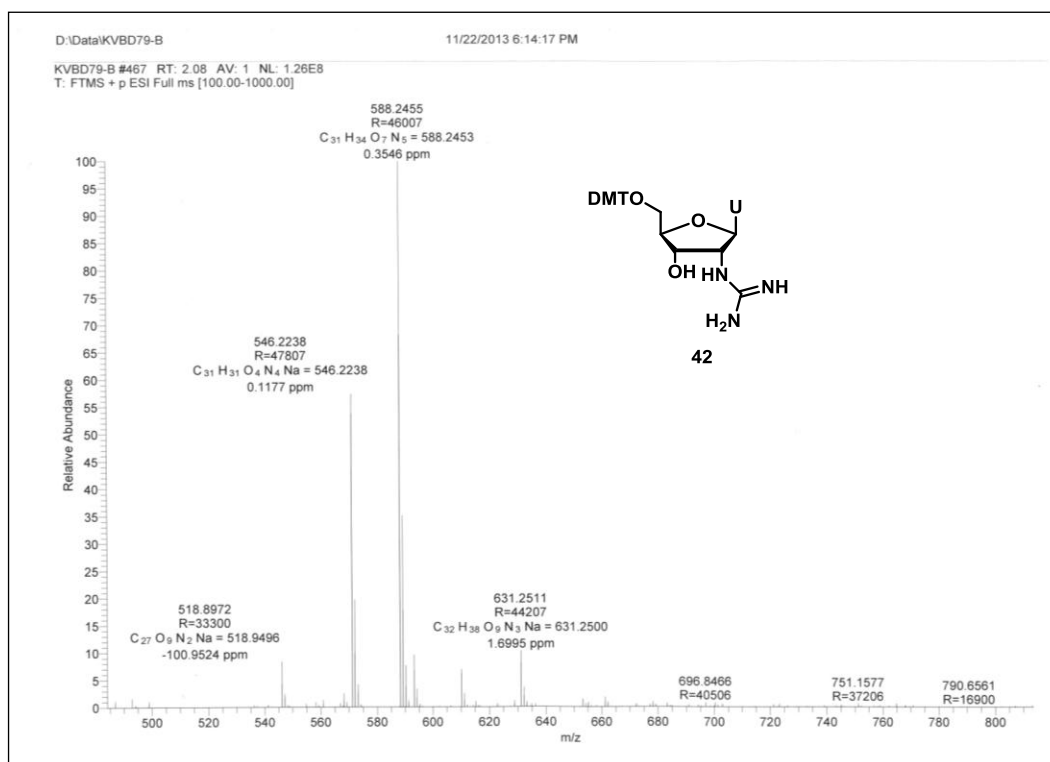
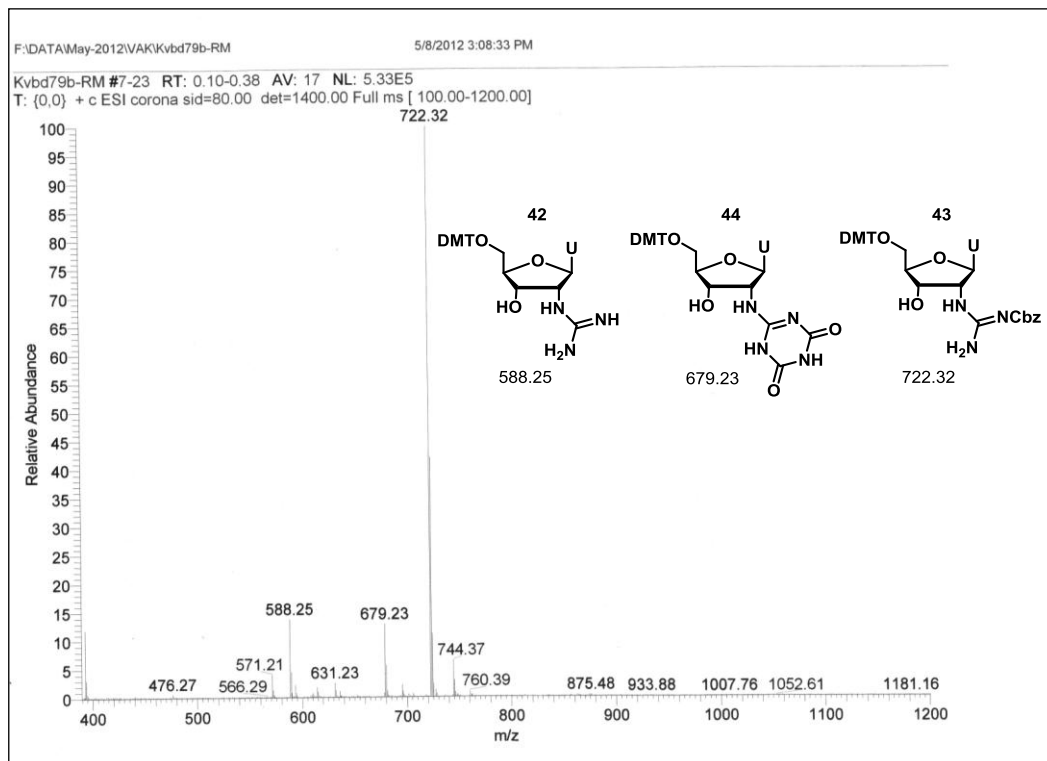
Mass (LCMS/HRMS) spectral data:



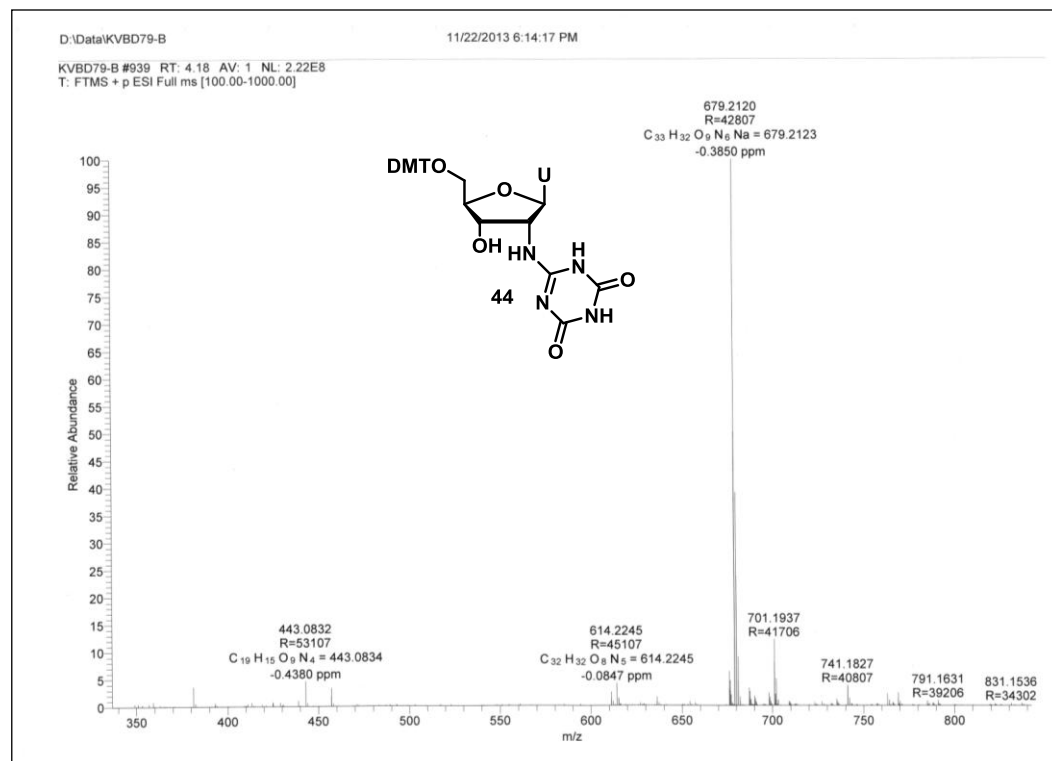
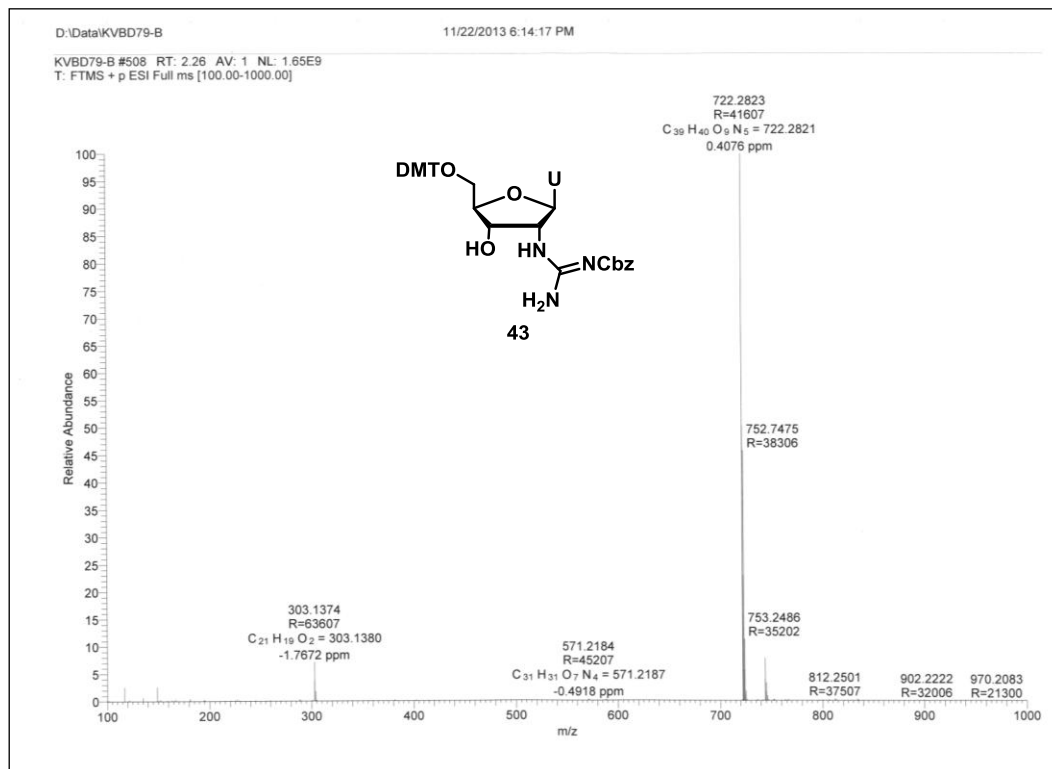
Chapter 2



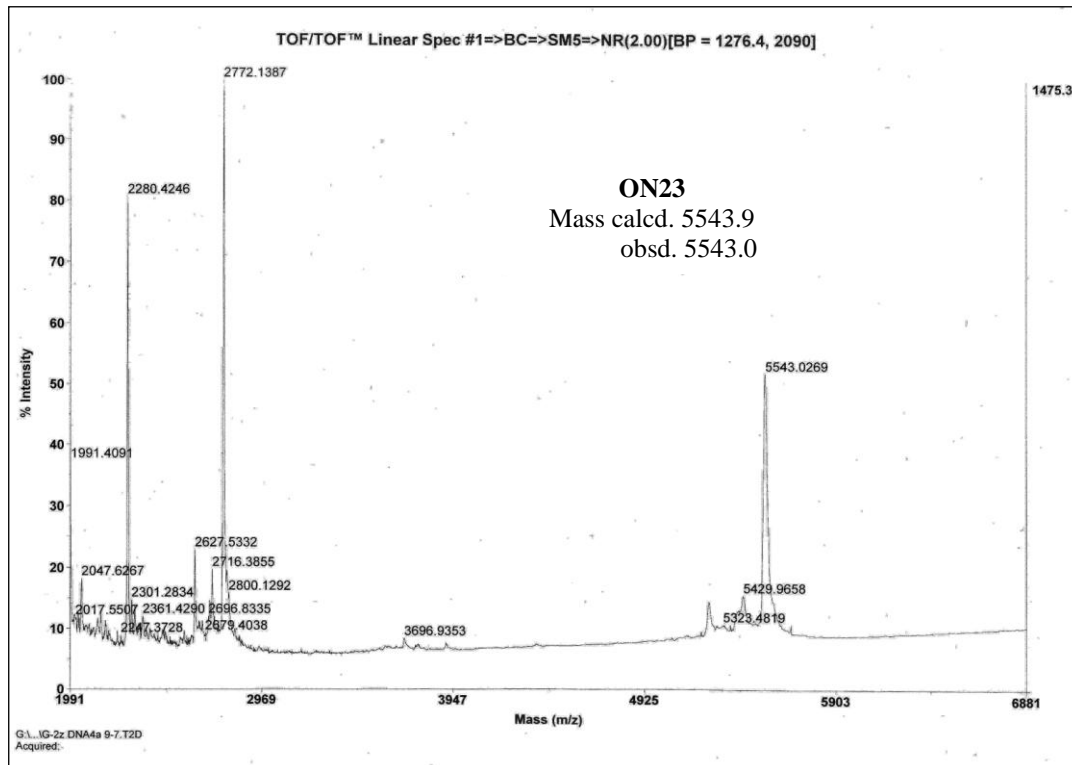
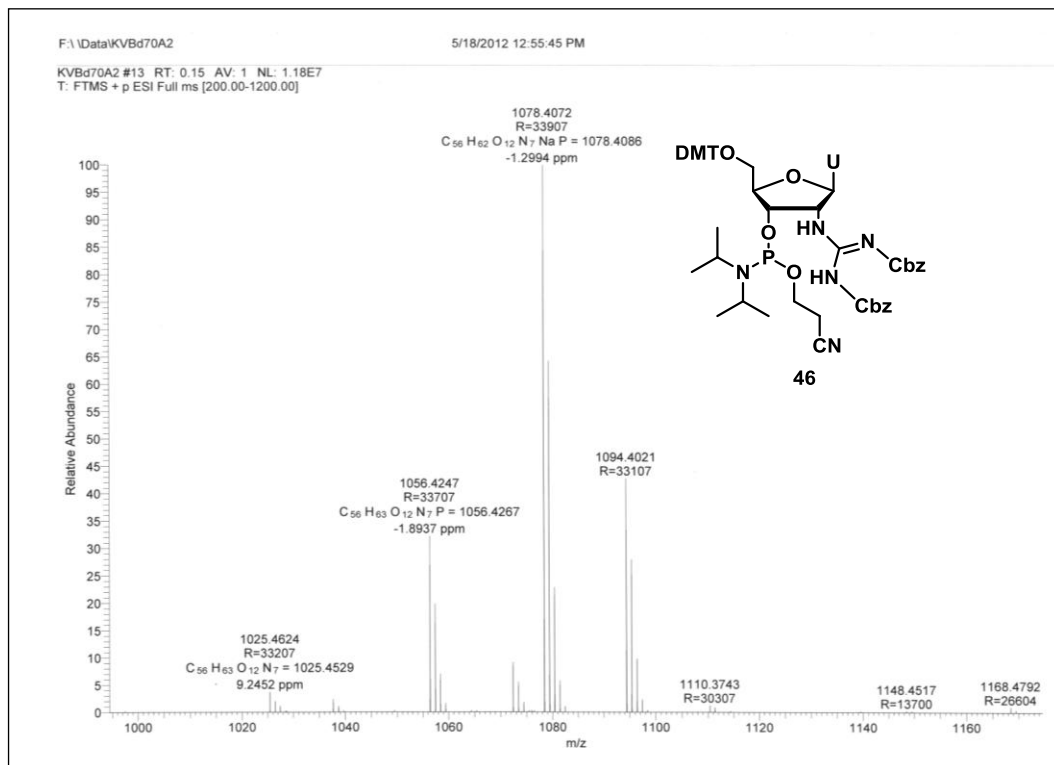
Chapter 2

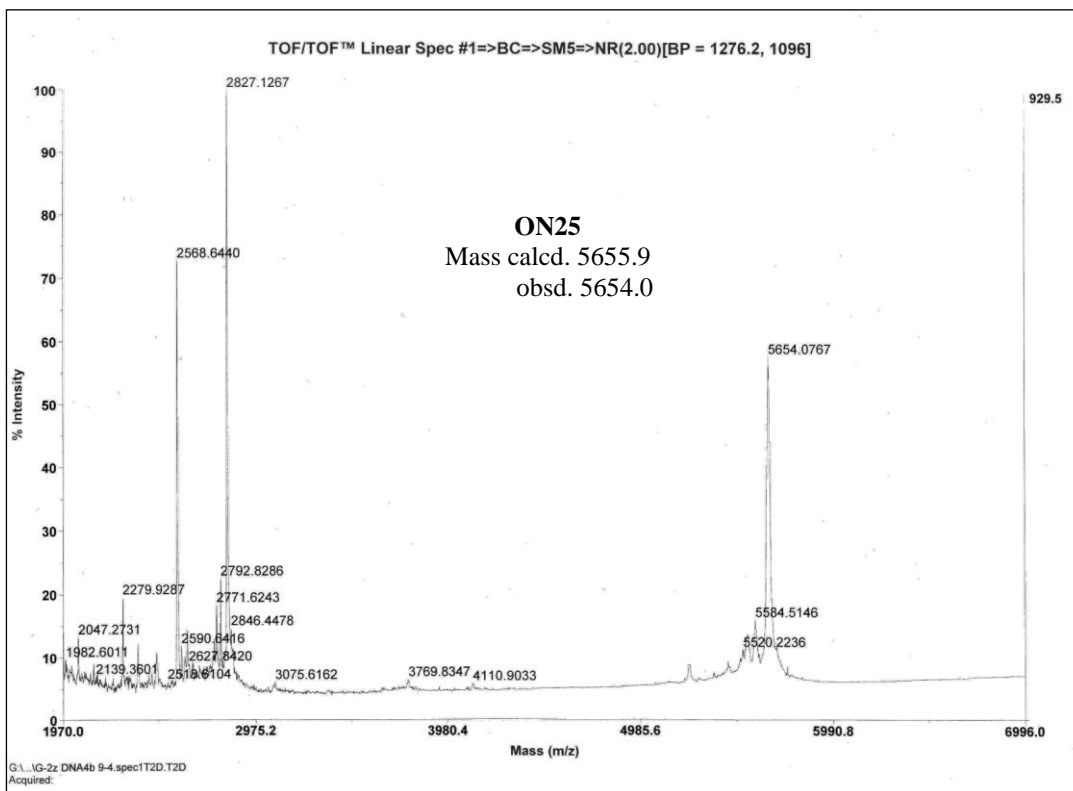
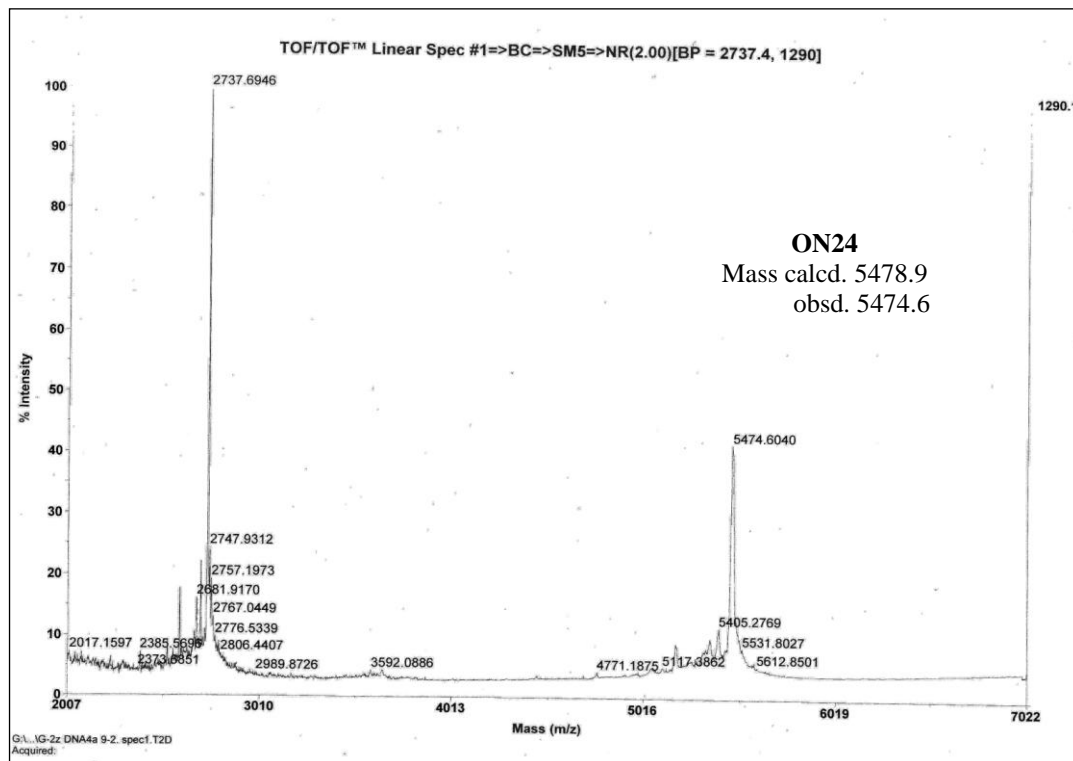


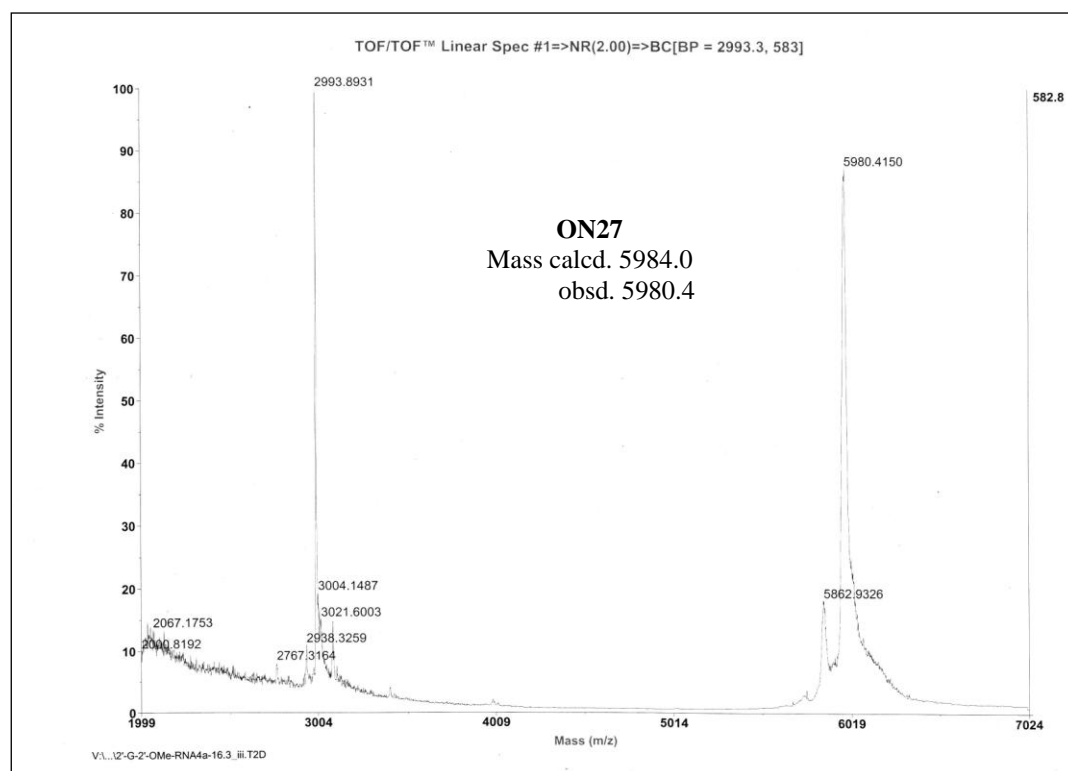
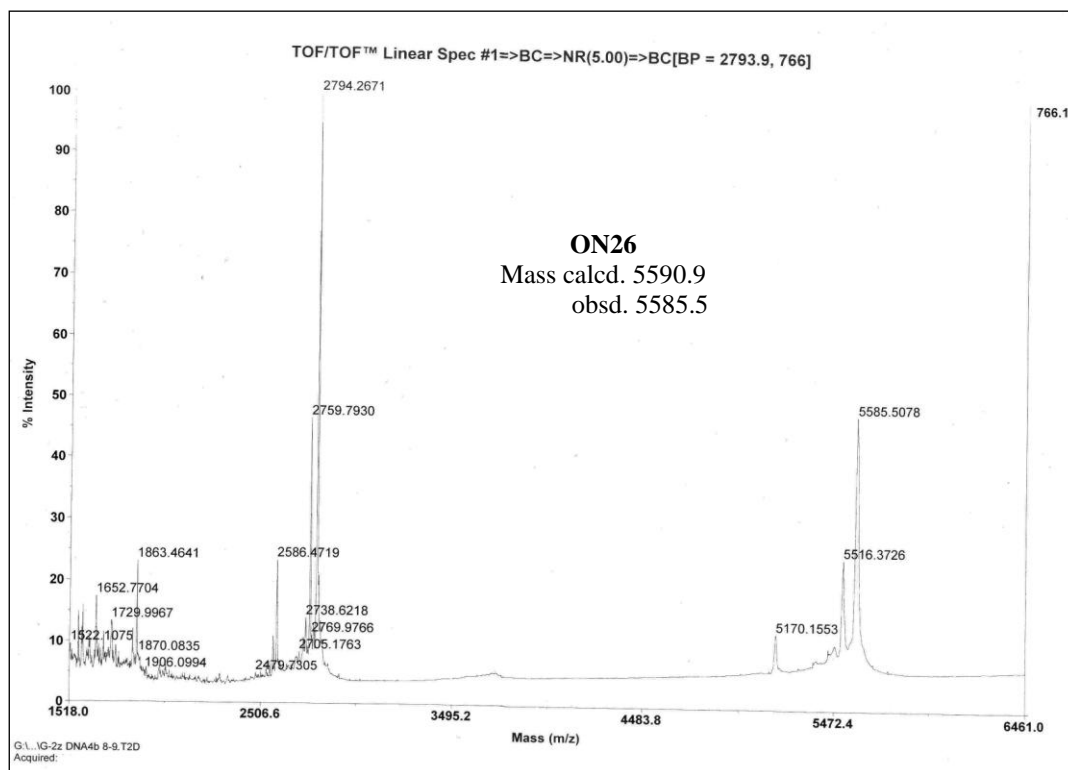
Chapter 2

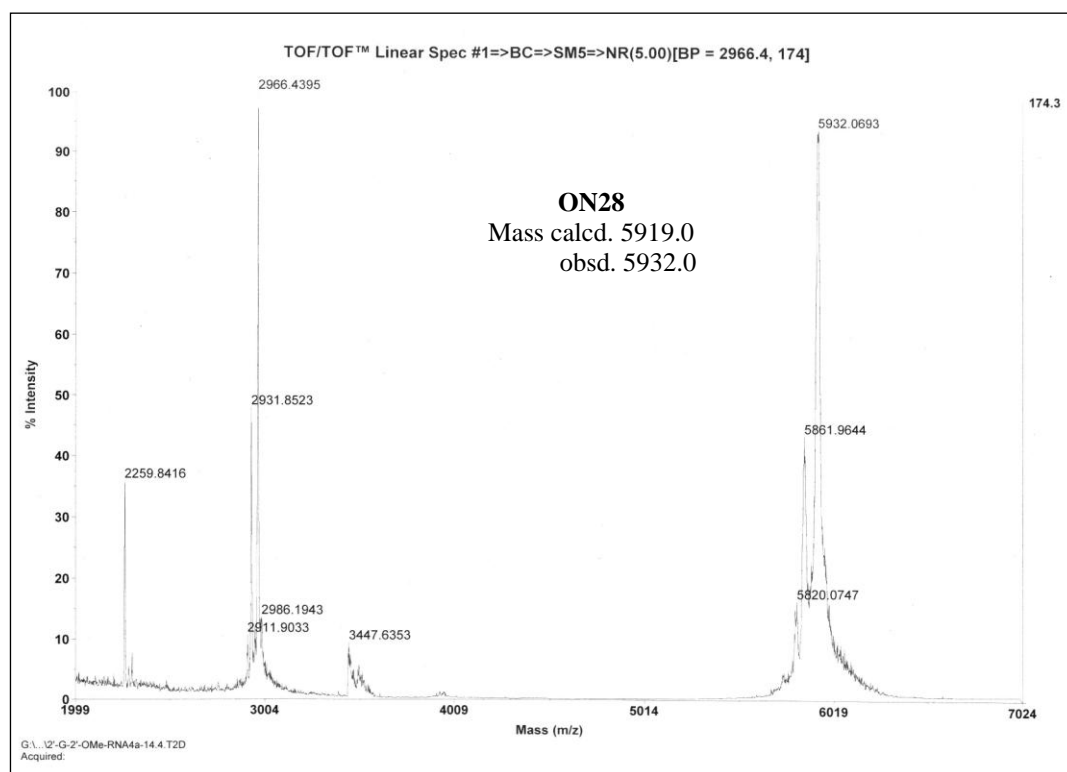


Chapter 2









2D References

1. (a)Bennett, C. F.; Swayze, E. E., *Annual review of pharmacology and toxicology* **2010**, 50, 259-293. (b)Kole, R.; Krainer, A. R.; Altman, S., *Nat. Rev. Drug discovery* **2012**, 11, 125-140.
2. (a)Martin, P., *Helv. Chim. Acta* **1995**, 78, 486-504. (b)Prakash, T. P., *Chemistry & Biodiversity* **2011**, 8, 1616-1641. (c)Freier, S. M.; Altmann, K.-H., *Nucleic Acids Res.* **1997**, 25, 4429-4443.
3. (a)Manoharan, M.; Akinc, A.; Pandey, R. K.; Qin, J.; Hadwiger, P.; John, M.; Mills, K.; Charisse, K.; Maier, M. A.; Nechev, L.; Greene, E. M.; Pallan, P. S.; Rozners, E.; Rajeev, K. G.; Egli, M., *Angew. Chem. Int. Ed.* **2011**, 50, 2284-2288. (b)Lima, F.; Prakash, T. P.; Murray, H. M.; Kinberger, G. A.; Li, W.; Chappell, A. E.; Li, C. S.; Murray, S. F.; Gaus, H.; Seth, P. P.; Swayze, E. E.; Crooke, S. T., *Cell* **2012**, 150, 883-894
4. (a)Teplova, M.; Minasov, G.; Tereshko, V.; Inamati, G. B.; Cook, P. D.; Manoharan, M.; Egli, M., *Nat. Struct. Biol.* **1999**, 6, 535-539. (b)Tereshko, V.; Portmann, S.; Tay, E. C.; Martin, P.; Natt, F.; Altmann, K.-H.; Egli, M., *Biochemistry* **1998**, 37, 10626–10634.
5. Guterstam, P.; Lindgren, M.; Johansson, H.; Tedebark, U.; Wengel, J.; Andaloussi, S.; Langel, U., *Biochem. J.* **2008**, 412, 307–313.
6. Kaur, H.; Babu, B. R.; Maiti, S., *Chem. Rev.* **2007**, 107, 4672-4697.
7. Teplova, M.; Wallace, S. T.; Minasov, G.; Tereshko, V.; Symons, A.; Cook, P. D.; Manoharan, M.; Egli, M., *Proc. Natl. Acad. Sci.U.S.A* **1999**, 96, 14240-14245.
8. Jain, M.; Bruice, P.; Szabó, I.; Bruice, T., *Chem. Rev.* **2012**, 112, 1284-1309.
9. Lu, K.; Duan, Q.-P.; Ma, L.; Zhao, D.-X., *Bioconjugate Chem.* **2010**, 21, 187-202.
10. Griffey, R. H.; Monia, B. P.; Cummins, L. L.; Freier, S.; Greig, M. J.; Guinasso, C. J.; Lesnik, E.; Manalili, S. M.; Mohan, V.; Owens, S.; Ross, B. R.; Sasmor, H.; Wancewicz, E.; Weiler, K.; Wheeler, P. D.; Cook, P. D., *J. Med. Chem.* **1996**, 39, 5100-5109.

Chapter 2

11. Sabahi, A.; Guidry, J.; Inamati, G. B.; Manoharan, M.; Wittung-Stafshede, P., *Nucleic Acids Res.* **2001**, 29, 2163-2170.
12. (a)Lesnik, E. A.; Guinosso, C. J.; Kawasaki, A. M.; Sasmor, H.; Zounes, M.; Cummins, L. L.; Ecker, D. J.; Cook, P. D.; Freier, S. M., *Biochemistry* **1993**, 32, 7832-7838. (b)Egli, M.; Minasov, G.; Tereshko, V.; Pallan, P. S.; Teplova, M.; Inamati, G. B.; Lesnik, E. A.; Owens, S. R.; Ross, B. S.; Prakash, T. P.; Manoharan, M., *Biochemistry* **2005**, 44, 9045-9057.
13. Gopalakrishna, V.; Kumar, V. A.; Ganesh, K. N., *Nucleosides and Nucleotides* **1992**, 11, 1263-1273.
14. Saneyoshi, H.; Seio, K.; Sekine, M., *J. Org. Chem.* **2005**, 70, 10453-10460.
15. (a)Lou, C.; Xiao, Q.; Brennan, L.; Light, M. E.; Vergara-Irigaray, N.; Atkinson, E. M.; Holden-Dye, L. M.; Fox, K. R.; Brown, T., *Bioorg. Med. Chem.* **2010**, 18, 6389-6397. (b)Seth, P. P.; Vasquez, G.; Allerson, C. A.; Berdeja, A.; Gaus, H.; Kinberger, G. A.; Prakash, T. P.; Migawa, M. T.; Bhat, B.; Swayze, E. E., *J. Org. Chem.* **2010**, 75, 1569-1581. (c)Cuenoud, B.; Casset, F.; Husken, D.; Natt, F.; Wolf, R. M.; Altmann, K.-H.; Martin, P.; Moser, H. E., *Angew. Chem. Int. Ed.* **1998**, 37, 1288-1291. (d)Buchini, S.; Leumann, C. J., *Tetrahedron Lett.* **2003**, 44, 5065-5068. (e)Prakash, T. P.; Manoharan, M.; Kawasaki, A. M.; Fraser, A. S.; Lesnik, E. A.; Sioufi, N.; Leeds, J. M.; Teplova, M.; Egli, M., *Biochemistry* **2002**, 41, 11642-11648.
16. Ross, B. S.; Springer, R. H.; Tortorici, Z.; Zimock, S., *Nucleosides and Nucleotides* **1997**, 16, 1641-1643.
17. Legorburu, U.; Reese, C. B.; Song, Q., *Tetrahedron* **1999**, 55, 5635-5640.
18. (a)Manoharan, M.; Prakash, T. P.; Barber-Peoc'h, I.; Bhat, B.; Vasquez, G.; Ross, B. S.; Cook, P. D., *J. Org. Chem.* **1999**, 64, 6468-6472. (b)Prakash, T. P.; Püschl, A.; Lesnik, E.; Mohan, V.; Tereshko, V.; Egli, M.; Manoharan, M., *Org. Lett.* **2004**, 6, 1971-1974.
19. Saneyoshi, H.; Okamoto, I.; Masaki, Y.; Ohkubo, A.; Seio, K.; Sekine, M., *Tetrahedron Lett.* **2007**, 48, 8554-8557.

Chapter 2

20. Kotikam, V.; Kumar, V. A., *Tetrahedron* **2013**, 69, 6404-6408.
21. (a)Jin, Y. H.; Liu, P.; Wang, J.; Baker, R.; Huggins, J.; Chu, C. K., *J. Org. Chem.* **2003**, 68, 9012-9018. (b)Divakar, K. J.; Reese, C. B., *J. Chem. Soc., Perkin Trans. 1* **1982**, 1171-1176.
22. (a)Prhavic, M.; Prakash, T. P.; Minasov, G.; Cook, P. D.; Egli, M.; Manoharan, M., *Org. Lett.* **2003**, 5, 2017-2020. (b)Blommers, M. J. J.; Natt, F.; Jahnke, W.; Cuenoud, B., *Biochemistry* **1998**, 37, 17714-17725.
23. Sinha, N. D.; Biernat, J.; McManus, J.; Koster, H., *Nucleic Acids Res.* **1984**, 12, 4539-4557.
24. Kang, S.-H.; Cho, M.-J.; Kole, R., *Biochemistry* **1998**, 37, 6235-6239.
25. (a)Honcharenko, D.; Barman, J.; Varghese, O. P.; Chattopadhyaya, J., *Biochemistry* **2007**, 46, 5635-5646. (b)Prakash, T. P.; Manoharan, M.; Fraser, A. S.; Kawasaki, A. M.; Lesnik, E. A.; Owens, S. R., *Tetrahedron Lett.* **2000**, 41, 4855-4859.
26. Yu, D.; Kandimalla, E. R.; Roskey, A.; Zhao, Q.; Chen, L.; Chen, J.; S., A., *Bioorg. Med. Chem.* **2000**, 8, 275-284.
27. (a)Altona, C.; Sundaralingam, M., *J. Am. Chem. Soc.* **1972**, 94, 8205-8212. (b)Obika, S.; Morio, K.-i.; Nanbu, D.; Imanishi, T., *Chem. Commun.* **1997**, 0, 1643-1644.
28. Zamecnik, P. C.; Stephenson, M. L., *Proc. Natl. Acad. Sci. USA* **1978**, 75, 280-284.
29. Agarwal, K. L.; Khorana, H. G., *J. Am. Chem. Soc.* **1972**, 94, 3578-3585.
30. Tennant, R. W.; Farrelly, J. G.; Ihle, J. N.; Pal, B. C.; Kenney, F. T.; Brown, A., *J. Virol.* **1973**, 12, 1216-1225.
31. (a)Lennox, K. A.; Owczarzy, R.; Thomas, D. M.; Walder, J. A.; Behlke, M. A., *Molecular Therapy—Nucleic Acids* **2013**, 2, e117. (b)Blechinger, J.; Pieper, H.; Marzenell, P.; Kovbasyuk, L.; Serva, A.; Starkuviene, V.; Erfleb, H.; Mokhir, A., *Chem. Commun.* **2013**, 49, 7397-7399.
32. Kole, R.; Krainer, A. R.; Altman, S., *Nature Rev. Drug Discov.* **2012**, 11, 125-140.
33. Nielsen, P. E.; Egholm, M.; Berg, R. H.; Buchardt, O., *Science* **1991**, 254, 1497-1500.

34. Summerton, J.; Weller, D., *Antisense Nucleic Acid Drug Dev.* **1997**, 7, 187-195.
35. Williams, D. M.; Benseler, F.; Eckstein, F., *Biochemistry* **1991**, 30, 4001-4009.
36. Damha, M. J.; Wilds, C. J.; Noronha, A.; Brukner, I.; Borkow, G.; Arion, D.; Parniak, M. A., *J. Am. Chem. Soc.* **1998**, 120, 12976–12977.
37. Koshkin, A. A.; Singh, S. K.; Nielsen, P.; Rajwanshi, V. K.; Kumar, R.; Meldgaard, M.; Olsen, C. E.; Wengel, J., *Tetrahedron* **1998**, 54, 3607-3630.
38. Watts, J. K.; Corey, D. R., *J. Pathol.* **2012**, 226, 365-379.
39. (a)Tang, J. Y.; Tamsamani, J.; Agrawal, S., *Nucleic Acids Res.* **1993**, 21, 2729–2735. (b)Lennox, K. A.; Behlke, M. A., *Pharm Res.* **2010**, 27, 1788–1799.
40. (a)Krutzfeldt, J.; Rajewsky, N.; Braich, R.; Rajeev, K. G.; Tuschl, T.; Manoharan, M.; Stoffel, M., *Nature* **2005**, 438, 685-689. (b)Krutzfeldt, J.; Kuwajima, S.; Braich, R.; Rajeev, K. G.; Pena, J.; Tuschl, T.; Manoharan, M.; Stoffel, M., *Nucleic Acids Res.* **2007**, 35, 2885–2892.
41. Bennett, C. F.; Swayze, E. E., *Annu. Rev. Pharmacol. Toxicol.* **2010**, 50, 259-293.
42. (a)Geary, R. S.; Yu, R. Z.; Levin, A. A., *Curr. Opin. Invest Drugs* **2001**, 2, 562-573. (b)Wu, H.; Lima, W. F.; Zhang, H.; Fan, A.; Sun, H.; Crooke, S. T., *J. Biol. Chem.* **2004**, 279, 17181–17189.
43. (a)Deleavey, G. F.; Watts, J. K.; Alain, T.; Robert, F.; Kalota, A.; Aishwarya, V.; Pelletier, J.; Gewirt, A. M.; Sonenberg, N.; Damha, M. J., *Nucleic Acids Res.* **2010**, 38, 4547–4557. (b)Gore, K. R.; Nawale, G. N.; Harikrishna, S.; Chittoor, V. G.; Pandey, S. K.; Höbartner, C.; Patankar, S.; Pradeepkumar, P. I., *J. Org. Chem.* **2012**, 77, 3233-3245.
44. Fabani, M. M.; Gait, M. J., *RNA* **2008**, 14, 336-346.
45. Bauman, J.; Jearawiriyapaisarn, N.; Kole, R., *Oligonucleotides* **2009**, 19, 1-13.
46. Bhan, P.; Bhan, A.; Hong, M.; Hartwell, J. G.; Saunders, J. M.; Hoke, G. D., *Nucleic Acids Res.* **1997**, 25, 3310–3317.
47. Guterstam, P.; Lindgren, M.; Johansson, H.; Tedebark, U.; Wengel, J.; Andaloussi, S. E.; Langel, U., *Biochem. J.* **2008**, 412, 307–313.

Chapter 2

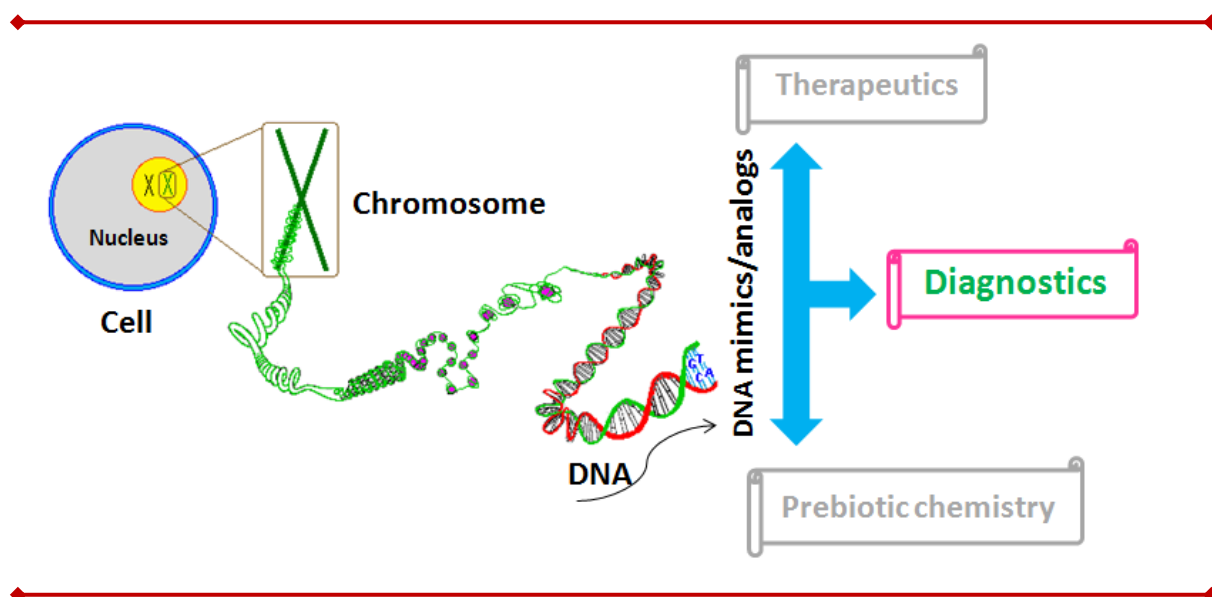
48. Thayer, J. R.; Wu, Y.; Hansen, E.; Angelino, M. D.; Rao, S., *J. Chromatogr. A* **2011**, 1218, 802-808.
49. Hari, Y.; Morikawa, T.; Osawa, T.; Obika, S., *Org. Lett.* **2013**, 14, 3702-3705.
50. Seela, F.; Kaiser, K., *Nucleic Acids Res.* **1987**, 15, 3113-3129.
51. Sugimoto, N.; Nakano, M.; Nakano, S., *Biochemistry* **2000**, 39, 11270-11281.
52. Dominski, Z.; Kole, R., *Proc. Natl. Acad. Sci. USA* **1993**, 90, 8673-8677.
53. Arzumanov, A.; Walsh, A. P.; Rajwanshi, V. K.; Kumar, R.; Wengel, J.; Gait, M. J., *Biochemistry* **2001**, 40, 14645-14654.
54. Kumar, R.; Singh, S. K.; Koshkin, A. A.; Rajwanshi, V. K.; Meldgaard, M.; Wengel, J., *Bioorg. Med. Chem. Lett.* **1998**, 8, 2219-2222.
55. Ivanova, G. D.; Arzumanov, A.; Abes, R.; Yin, H.; Wood, M. J.; Lebleu, B.; Gait, M. J., *Nucleic Acids Res.* **2008**, 36, 6418-6428.
56. Kurreck, J., *Eur. J. Biochem.* **2003**, 270, 1628-1644.
57. Nielsen, P. E.; Egholm, M.; Berg, R. H.; Buchardt, O., *Science* **1991**, 254, 1497-1500.
58. Madhuri, V.; Kumar, V. A., *Org. Biomol. Chem.* **2010**, 8, 3734-3741.
59. (a) Michel, T.; Debart, F.; Vasseur, J.-J., *Tetrahedron Letters* **2003**, 44, 6579-6582.
(b) Deglane, G.; Abes, S.; Michel, T.; Prévot, P.; Vives, E.; Debart, F.; Barvik, I.; Lebleu, B.; Vasseur, J.-J., *ChemBioChem* **2006**, 7, 684-692.
60. (a) Maier, M.; Barber-Peoc'h, I.; Manoharan, M., *Tetrahedron Letters* **2002**, 43, 7613-7616. (b) Prakash, T. P.; Puschl, A.; Lesnik, E.; Mohan, V.; Tereshko, V.; Egli, M.; Manoharan, M., *Organic Letters* **2004**, 6, 1971-1974.
61. (a) Robles, J.; Grandas, A.; Pedroso, E., *Tetrahedron* **2001**, 57, 179-194.
(b) Doronina, S. O.; Behr, J.-P., *Tetrahedron Letters* **1998**, 39, 547-550.
62. Stanzl, E. Z.; Trantow, B. M.; Vargas, J. R.; Wender, P. A., *Acc. Chem. Res.* **2013**, DOI: 10.1021/ar4000554.
63. Dempcy, R. O.; Browne, K. A.; Bruice, T. C., *J. Am. Chem. Soc* **1995**, 117, 6140-6141.

Chapter 2

64. Zhou, P.; Wang, M.; Du, L.; Fisher, G. W.; Waggoner, A.; Ly, D. H., *J. Am. Chem. Soc* **2003**, 125, 6878-6879.
65. Gokhale, S. S.; Kumar, V. A., *Org. Biomol. Chem.* **2010**, 8, 3742-3750.
66. Shrestha, A. R.; Kotobuki, Y.; Hari, Y.; Obika, S., *Chem. Comm.* **2013**, DOI: 10.1039/C3CC46017G.
67. McGee, D. P. C.; Vaughn-Settle, A.; Vargeese, C.; Zhai, Y., *J. Org. Chem.* **1996**, 61, 781-785.
68. Prakash, T. P.; Pusehl, A.; Manoharan, M., *Nucleosides Nucleotides & Nucleic Acids* **2007**, 26, 149-159.
69. Manoharan, M.; Prakash, T. P.; Barber-Peoc'h, I.; Bhat, B.; Vasquez, G.; Ross, B. S.; Cook, P. D., *J. Org. Chem.* **1999**, 64, 6468-6472.
70. Sekine, M.; Oeda, Y.; Iijima, Y.; Taguchi, H.; Ohkubo, A.; Seio, K., *Org. Biomol. Chem.* **2011**, 9, 210-218.
71. Butora, G.; Kenski, D. M.; Cooper, A. J.; Fu, W. L.; Qi, N.; Li, J. J.; Flanagan, W. M.; Davies, I. W., *J. Am. Chem. Soc* **2011**, 133, 16766-16769.
72. (a)Kumar, P.; Madsen, C. S.; Nielsen, P., *Bioorg. Med. Chem. Lett.* **2013**, 23, 6847-6850. (b)Madsen, C. S.; Witzke, S.; Kumar, P.; Negi, K.; Sharma, P. K.; Petersen, M.; Nielsen, P., *Chem. Eur. J.* **2012**, 18, 7434-7442.

CHAPTER 3

Applications of synthetic DNA mimics in diagnostics



Section 3A: Design and synthesis of (*S*)-polycarbamate nucleic acids ((*S*)-PCNA) and their biophysical evaluation

3A.1 Introduction

A potential modified oligonucleotide (AON) for applications in biology or diagnostics should have enhanced binding affinity to complementary nucleic acids and increased biological stability. Amongst several backbone-modified AONs reported so far, those involving replacement of the phosphodiester group with neutral linkers, in the absence of inter-strand electrostatic repulsions with target nucleic acids show improvement in their binding affinity. Complete replacement of sugar-phosphate backbone of DNA by achiral, acyclic and uncharged scaffold in PNA is the most prominent example and has been extensively studied in the last two decades since its discovery in 1991 (Figure 1).¹ In PNA, nucleobases are attached to the polyamide backbone through conformationally rigid tertiary amide linker. The strong and sequence specific binding of PNA to both complementary DNA and RNA along with its unique strand invasion properties has proved to be extremely beneficial to both chemists and biologists working towards the development of potent diagnostic assays² and antisense oligonucleotides.^{1a, 3}

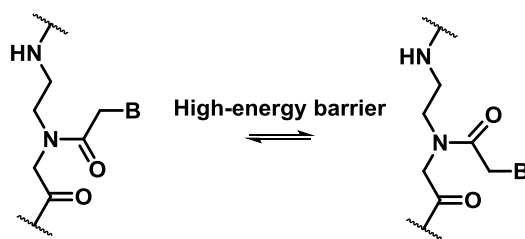


Figure 1 The case of *cis-trans* isomerization in PNA due to *tert*-amide linker.

Despite having several advantages such as resistance to cellular enzymes⁴ (proteases and nucleases) and strong binding to cDNA/RNA, absence of selectivity in binding in either parallel or antiparallel orientation to the cDNA/RNA is a limitation of achiral PNA in application perspective.^{1b, 5} Also PNA monomers are mixture of *cis/trans* isomers (Figure 1) due to the tertiary amide linker carrying nucleobases, with high energy barrier and are

not inter-convertible. At oligomer level, PNA becomes a mixture of *cis/trans* isomer at each monomer unit causing it to be a mixture of 2^n number of isomeric PNA. PNA:RNA/DNA duplex structure prefer the P-form, much wider and more slowly winding helix and differs considerably from A or B form structures adopted by RNA:RNA/DNA or DNA:DNA duplexes respectively.⁶ These studies imply that PNA is not a perfect structural mimic of DNA or RNA and enough scope exists for chemical modifications for further improvement. Thus, to design a PNA analogue as a better DNA mimic it is desirable to construct a backbone to overcome some of the drawbacks of PNA. Attempts have been made to introduce asymmetry into PNA backbone in such a way that it becomes chiral as well as induces certain conformational constraints towards orientation selective binding to complementary DNA/RNA sequences (as discussed in the Chapter 1).⁷

The modification of the DNA with the carbamate linkages preceded most other modifications since the discovery of antisense technology. This non-ionic modification of DNA involves replacement of phosphodiester linkage with carbamate linkage. The synthesis was achieved as early as in 1974 by Michael J. Gait *et al.*,⁸ where a dinucleotide analogue containing the oxyformamido-linkage, thymidinylformamido-[3'(O) 5'(C)]-5'-deoxythymidine was synthesized (Figure 2). These linkages are stable under physiological conditions and are resistant to nuclease action.^{8b}

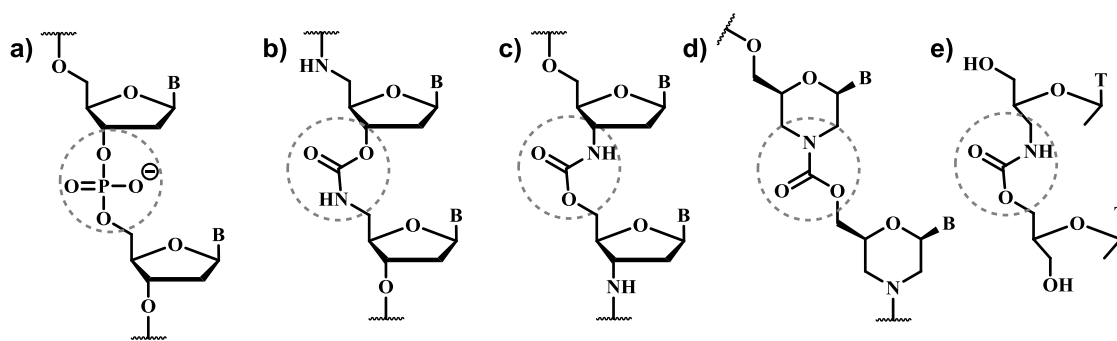


Figure 2 Chemical structures of DNA (a) and Carbamate-linked nucleic acid analogues: 3'-O-CO-NH-5' linked 2'-deoxyribose (b), 3'-NH-CO-O-5' linked 2'-deoxyribose (c), carbamate linked morpholine (d) and carbamate-linked acyclic sugar unit (e).

Though the thermal denaturation studies of hexadeoxythymidinyl carbamate-linked (3'-O-CO-NH-5', Figure 2b) oligomers were inconclusive due to the low thermal stability of short A:T duplexes,^{8c} the cytosine containing oligonucleosides with an identical backbone formed the stable duplexes with the complementary DNA.⁹ The chimeric oligonucleotides containing one, two or three internucleosidic carbamate units (-3'-NH-CO-O-5', Figure 2c) at 3'-end, were found to have increased nuclease resistance but had no significant influence on the duplex thermal stability.¹⁰ The carbamate linkages were extended to the morpholinocarbamates, where the carbamate linkage was introduced between the two morpholino units rather than between the two ribose sugars. The hexa morpholino-oligomer with the carbamate linkage (Figure 2d) carrying cytosine as the nucleobase formed much stable duplexes with the complementary DNA. The hypochromicity upon binding to DNA for these carbamate-linked oligomers was much higher than that observed for the control sequences.^{8d} In another report, chimeric carbamate-internucleoside dimer block of 2'-deoxy-2', 3'-secothymidine, (Figure 2e) protected in the 5'-position, was incorporated into the oligonucleotides by phosphoramidate chemistry. The release in the ring strain of ribose sugar had detrimental effect and these chimeric oligonucleotides showed less affinity towards complementary DNA compared to unmodified strand.¹¹

3A.2 Rationale and objectives of the present work

Our group earlier reported the synthesis of pyrrolidinyl carbamate oligonucleotides, but the flexibility in the linker group led to destabilization of the complexes.¹² This was then modified by introducing a fresh approach to make the DNA mimics with an oligocarbamate backbone in which the structure is closely related to PNA. In this context, our group recently reported the simple variation in the structure of PNA leading to *R*-polycarbamate nucleic acids (*R*-PCNA) which allowed the presence of a chiral center in the monomer unit and conserves six atom repeating backbone.¹³ The chiral centre corresponds to C-4' in natural nucleic acids and the individual nucleobases were attached through an amide linker as in PNA to the 2,3-diamino-1-propanol core. The linking of individual

monomers derived from L-serine, *via* carbamate linkage to form *R*-PCNA (Figure 3) conserved six atom sugar-phosphate or aminoethylglycyl repeating backbones of DNA or PNA respectively. The nucleobase attachment to the carbamate backbone *via* a secondary amide function could circumvent the likelihood of *cis/trans* isomers present in tertiary amide linker groups in PNA (Figure 1).^{1a, 14}

The additional hydrogen bonding sites in polycarbamate backbone rendered it more water solubility as compared with PNA. The designed polycarbamate backbone (PCNA, Figure 3) was chiral in nature as each monomer unit is chiral. The chirality in the system of PCNA therefore showed discrimination towards the parallel and antiparallel mode of binding and also discrimination while binding with DNA or RNA. The hydroxy group of each monomer was activated as a *p*-nitrophenyl carbonate and the oligomerization was achieved by solid phase carbamate synthesis.

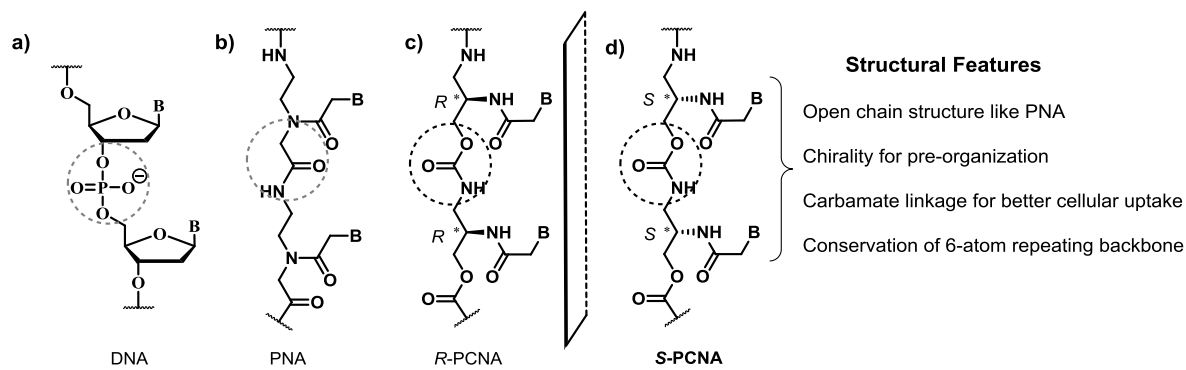


Figure 3 Chemical structures of DNA, PNA and our designed polycarbamate nucleic acids.

It was concluded from this work that *R*-PCNAs could discriminate between DNA and RNA, but preferred to bind to cDNA better than cRNA in an antiparallel orientation. In the present work, the change in chirality of *R*-PCNA backbone is undertaken. The synthesis of *S*-PCNA from D-serine was planned to investigate if this would alter the binding selectivity.¹⁵ To realize the synthetic goal, we conceived a desymmetrization-like strategy to allow the synthesis of both *S*-PCNA and *R*-PCNA from naturally occurring L-serine.

3A.3 Retrosynthetic analysis

Natural amino acid, L-serine, is the ideal chemical source to achieve both the enantiomeric oligocarbamates *R*-PCNA and *S*-PCNA. The retrosynthetic disconnections for *S*-PCNA are similar to the reported *R*-PCNA from our group,¹³ the only change being in the order of the chemical reactions performed, starting from natural amino acid L-serine. The reduction of the carboxylic acid functionality in L-serine would furnish the symmetrical serinol unit and one can break this symmetry by the simple change in the protection/deprotection sequence to arrive at both the enantiomers. In the case of *R*-PCNA, the hydroxy functional group originally present in the L-serine (blue colored, Figure 4) was involved in as the activated carbonate. On the other hand, the alcohol resulting from reduction of carboxylic acid functionality in L-serine (pink colored, Figure 4), when involved in the carbonate activation, would lead to *S*-PCNA.

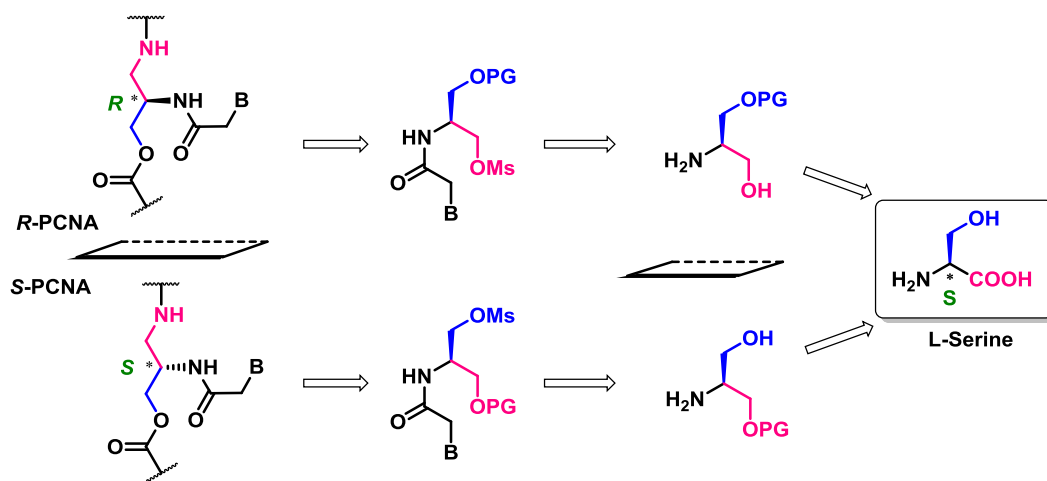


Figure 4 Retrosynthetic analysis of *R*-PCNA and *S*-PCNA from single chiral source, L-serine.

3A.4 Synthesis of PCNA oligomers

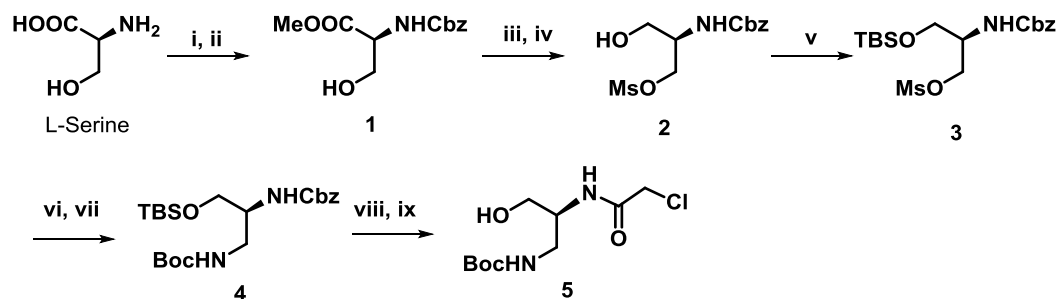
3A.4.1 Synthesis of activated carbonate monomers

Our synthetic route starts with the Cbz protection of amine functionality in the natural amino acid L-Serine followed by the esterification using thionyl chloride in MeOH (Scheme 1). The mesylation of free hydroxyl group was performed at 0 °C in pyridine for 10 min. The crude mesylate ester was subjected to reduction with sodium borohydride to

Chapter 3

furnish alcohol **2** which was protected as a *tert*-butyldimethylsilyl ether, to get **3**. SN^2 displacement of mesylate by azide and subsequent reduction of azide to amino functionality using Raney-Ni followed by protection with $(\text{Boc})_2\text{O}$ gave compound **4**. The Cbz group was deprotected under hydrogenation conditions using Pd/ H_2 and the resulted amine was chloroacetylated using chloroacetyl chloride and K_2CO_3 to yield **5**. During the chloroacetylation of the free amine, which usually completes in ~ 90 min, the *-O*-TBS also got deprotected (under basic conditions) by keeping the reaction for longer times furnishing the chloro alcohol **5**.

Scheme 1 Synthesis of precursors for Boc-strategy

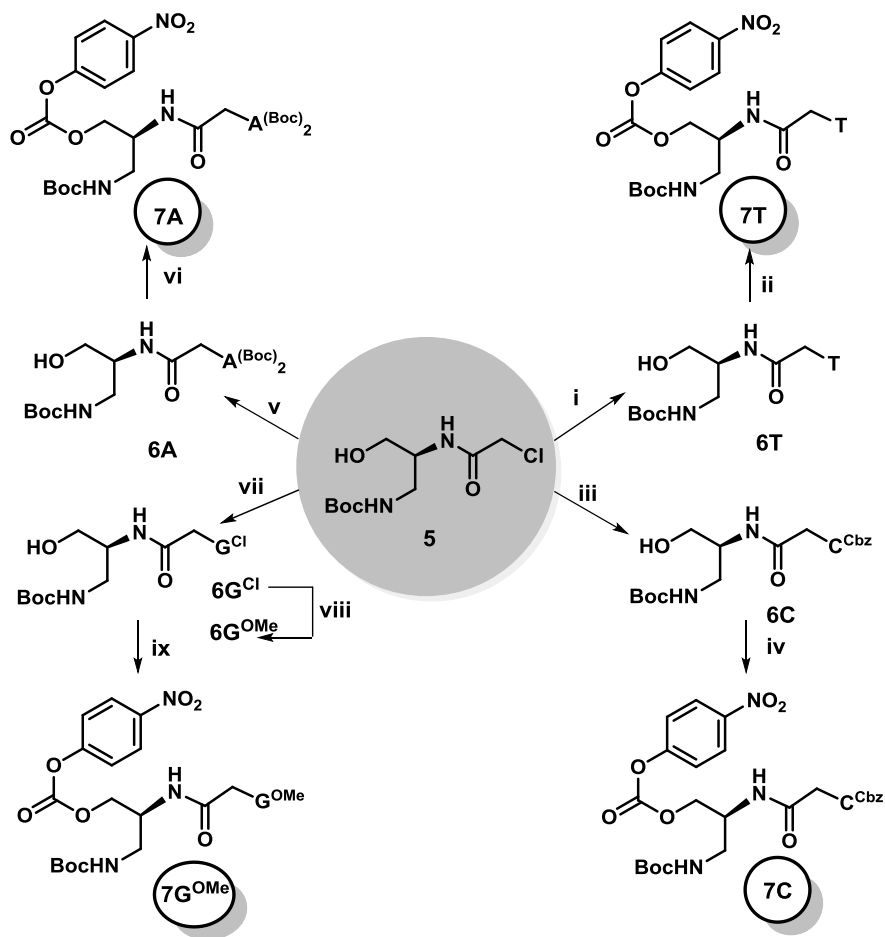


Reagents and conditions: (i) Cbz-Cl, NaHCO_3 , H_2O , 12 h, 70% (ii) SOCl_2 , Et_3N , MeOH, 6 h, 90% (iii) Ms-Cl, Py, 0 °C, dry DCM, 15 min (iv) NaBH_4 , MeOH, 0 °C-rt, 75% over two steps (v) TBS-Cl, imidazole, dry DCM, 5 h, 95% (vi) NaN_3 , dry DMF, 60 °C, 76% (vii) Raney Ni, H_2 , 40 psi, $(\text{Boc})_2\text{O}$, EtOAc, 4 h, 84% (viii) H_2 , Pd/C MeOH, 50 psi, 6 h (ix) ClCOCH_2Cl , K_2CO_3 , Dioxane: H_2O (1:1), 8 h, 80% over two steps.

The genetic nucleobases thymine (T), cytosine (C), adenine (A) and guanine (G) with suitable protection were successfully reacted with the intermediate **5**, through the SN^2 displacement of chloride with the nucleobases as the incoming nucleophiles to furnish the precursors **6T**, **6C**, **6A** and **6G^{Cl}** for solid phase carbamate synthesis. The basic hydrolysis of 2-amino-6-chloropurine derivative (**6G^{Cl}**) to guanine derivative (**6G^{OMe}**) (using 0.7M NaOH/MeOH) resulted the *-O*-methylated guanine rather than the free guanine derivative. The *O*-methylated guanine derivative can be hydrolysed to guanine derivative under the strong acidic conditions (final cleavage conditions of oligo from resin), as observed in the

case of *R*-PCNA, the *O*-methylated guanine itself was used as a precursor for the *S*-PCNA solid phase carbamate synthesis.

Scheme 2 Synthesis of carbonate monomers carrying genetic alphabets T, C, A and G



Reagents and conditions: (i) Thymine, K₂CO₃, dry DMF, rt, overnight, 91% (ii) *p*-NO₂-PhOCOCl, dry DCM, pyridine, 70% (iii) Cbz-cytosine, K₂CO₃, dry DMF, overnight, 80% (iv) *p*-NO₂-PhOCOCl, dry DCM, pyridine, 72% (v) (Boc)₂-Adenine, K₂CO₃, dry DMF, overnight, 75% (vi) *p*-NO₂-PhOCOCl, dry DCM, pyridine, 86% (vii) 2-amino-6-chloropurine, K₂CO₃, dry DMF, overnight, 80% (viii) 0.7 M NaOH, MeOH, rt, 10 h, 85% (ix) *p*-NO₂-PhOCOCl, dry DCM, pyridine, 66%.

In the case of *R*-PCNA monomer, the chirality was derived from L-serine.¹³ To synthesize its enantiomeric *S*-PCNA, in the present work also, we used L-serine as a chiral source. The configurational change was simply achieved through the change in the order of

the protection/deprotection reactions performed. The observed equal and opposite rotation values of the intermediates obtained (Figure 5) substantiate their opposite configurational identity.

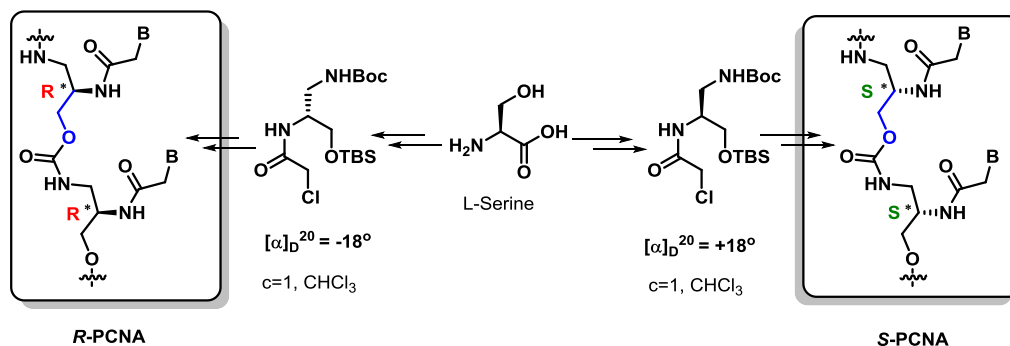


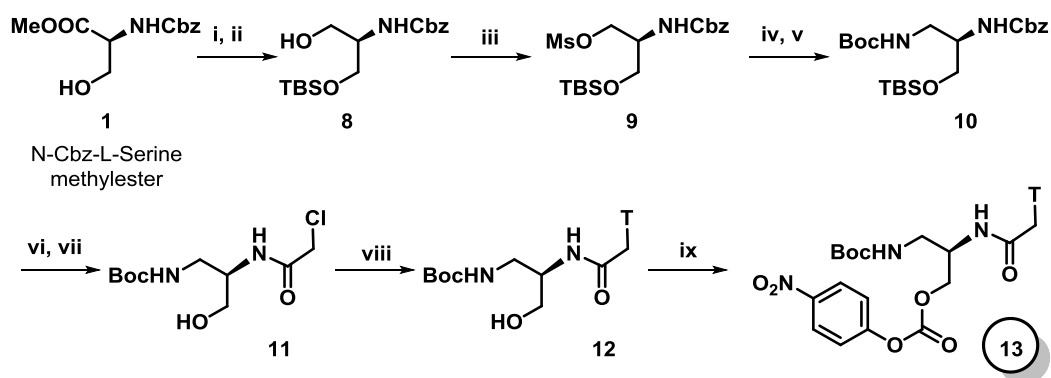
Figure 5 Specific rotation of the key intermediates from L-serine en route to the enantiomeric *R/S*-PCNA.

To obtain the desired carbamate linkages, towards oligocarbamate synthesis, we used the same protocol as that reported for *R*-PCNA synthesis. The precursors containing free hydroxyl group in **6T**, **6C**, **6A** and **6G^{OMe}** were treated with *p*-nitrochloroformate to form the activated carbonate monomers **7T**, **7C**, **7A** and **7G^{OMe}** respectively.

3A.4.2 Synthesis of *R*-PCNA thymine monomer

The free hydroxy group in protected L-serine derivative **1** was converted to TBS ether, followed by the reduction of the ester group using NaBH₄ in MeOH gave alcohol **8**, Scheme 3. The alcohol **8** was then converted into azide *via* mesylate **9** followed by reduction over Raney–Ni/H₂ and *in situ* -Boc protection of the primary amino group to get **10**. Compound **10** was subjected to hydrogenation over 10% Pd–C to remove the -Cbz protection and the free amino group was acetylated using chloroacetyl chloride and K₂CO₃ in dioxane–water, furnished the desilylated (as discussed in *S*-PCNA synthesis, Scheme 3) product **11**. The nucleobase thymine was reacted with **11** in the presence of K₂CO₃ to get a precursor **12**, which was further activated as carbonate **13**, prone to nucleophilic attack by amines and could lead to the carbamate linkage.

Scheme 3 Synthesis of T-monomer for R-PCNA synthesis via Boc-strategy



Reagents and conditions (i) TBS-Cl, imidazole, dry DCM, 5 h, 95% (ii) NaBH₄, MeOH, 0 °C-rt, 81% over two steps (iii) Ms-Cl, Py, 0 °C, dry DCM, 15 min (iv) NaN₃, dry DMF, 60 °C, 76% (v) Raney Ni, H₂, 40 psi, (Boc)₂O, EtOAc, 4 h, 87% (vi) H₂, Pd/C MeOH, 50 psi, 6 h (vii) ClCOCH₂Cl, K₂CO₃, Dioxane:H₂O (1:1), 8 h, 80% over two steps (viii) Thymine, K₂CO₃, dry DMF, rt, overnight, 88% (ix) *p*-NO₂-PhOC(=O)Cl, dry DCM, pyridine, 72%.

3A.4.3 Activated carbonates to carbamate backbone nucleic acids

Alcohols can be activated as carbonates or carbamates and which would be very reactive towards nucleophiles, Figure 6. In the present work the alcohol was activated as carbonate monomer (**7T/7C/7A/7G^{OMe}**), which on treatment with the free amino compound would displace the *p*-nitrophenol, a good leaving group, to furnish the desired carbamate linkage between the two monomers, in a growing oligocarbamate chain.

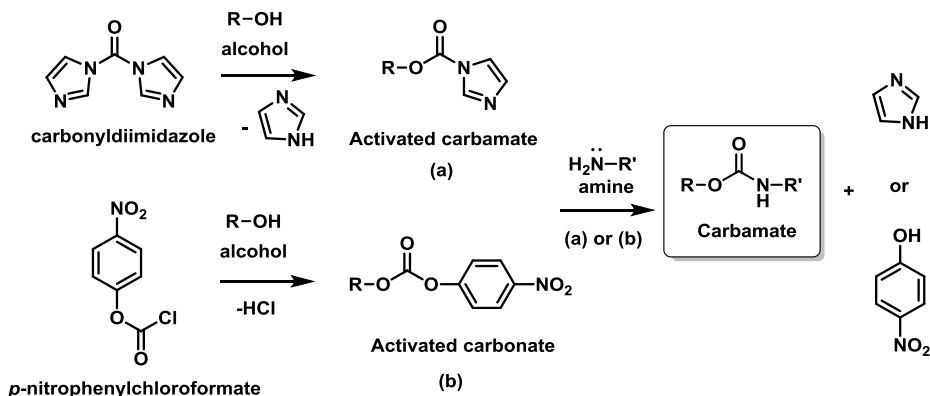


Figure 6 Formation of carbamate linkage from the reaction of activated carbonate/carbamate with alkyl amine.

The advantage of solid phase synthesis of oligocarbamates over regular solid phase peptide synthesis is that the (a) oligocarbamate synthesis does not involve any expensive coupling reagents (b) excess unreacted monomers can be recovered as they are not contaminated with any other coupling reagents, contrary to the case of peptide synthesis.

3A.5 General principles of solid phase synthesis

In contrast to the solution phase method, the solid phase peptide synthesis strategy devised by Merrifield,¹⁶ offers great advantages. In this method, the C-terminal amino acid is linked to polymeric beads having reactive functional groups, which also act as a permanent protection for the carboxylic acid. The next *N*^α-protected amino acid is coupled to the resin bound amino-functionality either by using an active pentafluorophenyl (pfp) or 3-hydroxy-2, 3-dihydro-4-oxo- benzotriazole (Dhbt) ester or by an *in situ* activation with carbodiimide reagents along with HOBt. These reagents are used in excess to drive the reaction to completion (>95%). The unreacted excess reagents are then washed out and the amino group deprotection, coupling reaction and washing steps are repeated until the desired peptide is achieved. The need to purify the coupling at every step is obviated. Finally, the resin bound peptide and the side chain protecting groups are cleaved in one step. The advantages of solid phase synthesis are (i) all the reactions are performed in a single vessel minimizing the loss due to transfer, (ii) large excess of activated monomer carboxylic acid component can be used resulting in high coupling efficiency, (iii) excess reagents can be removed by simple filtration and washing steps and (iv) the method is amenable to automation.

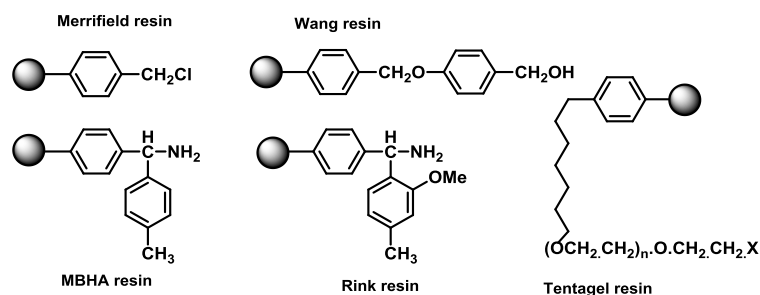


Figure 7 Few representative structures of different functional groups linked to resins used in SPPS.

3A.6 Solid phase synthesis of oligocarbamates

The solid phase synthesis of carbamate-linked nucleic acids^{18, 19} was carried out as shown in Figure 8 (synthesis achieved from C-terminus to N-terminus). The 4-methylbenzhydryl amine resin (MBHA resin) was first functionalized *via* amide linkage with the L-lysine, using HBTU/HOBt as coupling reagents and DIPEA as base with the loading value of 0.35 mmol/g and was confirmed by picric acid assay (Merrifield & Stewart, 1996; Erickson & Merrifield, 1976). L-lysine is known to increase the overall solubility of the uncharged oligomers. The α -amino group of the lysine anchored to the resin was freed by deprotecting the *t*-Boc group with 50% TFA in DCM followed by neutralization with 5% DIPEA in DCM. A mixture of 3-4 equivalents of activated monomer and catalytic amount of DIPEA in DMF was added to the resin having free amine reactive sites. The coupling reaction was completed in 2 h to get the first monomer attached to the lysine residue on resin through the carbamate linkage. The excess reagents were washed and recovered. For second monomer coupling, the Boc-NH- of the attached monomer was deprotected to get the TFA-salt of free amine which neutralized again with DIPEA, followed by the next monomer coupling. This deprotection-neutralization-coupling cycle was repeated till the oligomers of required sequence and length were obtained. The deprotection of the *N*-Boc protecting group and the coupling reaction were monitored by Kaiser's test.¹⁷

In the present work, the standard method of peptide-resin cleavage to the oligocarbamates was extended, the resin bound oligomers were subjected to acidic hydrolysis by treating with strong acids trifluoromethanesulphonic acid (TFMSA) in trifluoroacetic acid (TFA)¹⁸ to yield the carbamate oligomer with amide-terminus.

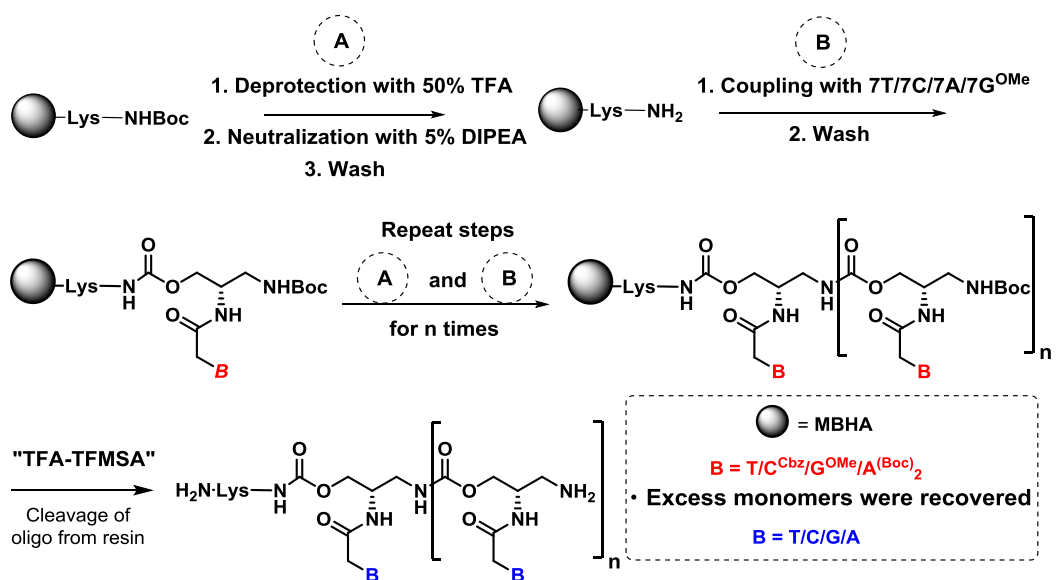


Figure 8 Schematic representation of solid phase synthetic cycle.

3A.7 Purification and characterization of the S-PCNA oligomers

Using the above mentioned protocol, we synthesized the octathyminyl carbamate sequence PCNA1, mixed pyrimidine sequence PCNA2 (part of *b2a2* gene) and purine-pyrimidine mixed sequence PCNA3 (part of micro-RNA). The purity of the cleaved crude oligomers was checked by analytical RP-HPLC (C18 column, CH₃CN-H₂O system containing 0.1% TFA), which indicated around 55-75% purity. These were subsequently purified by reverse phase HPLC on a semi preparative C18 column. The purity of the oligomers was again ascertained by analytical RP-HPLC and their integrity was confirmed by MALDI-TOF mass spectrometric analysis, Table 1.

The complementary DNA oligonucleotides DNA1-DNA3 and the control DNA sequences DNA4-DNA6 were synthesized on Mermade 4 DNA synthesizer using standard β -cyanoethyl phosphoramidate chemistry.¹⁸ The oligomers were synthesized from 3' to 5' direction on solid support (control pore glass (CPG) resin), followed by ammonia treatment. The oligonucleotides were purified by RP-HPLC on a C18 column and found to be more than 95% pure and were used without further purification in the biophysical studies of PCNAs.

Table 1 Synthesized oligomers and their MALDI-TOF analysis

Code	S-PCNA/DNA oligomers synthesized	Mass	
		Calcd.	Obsd.
S-PCNA1	H'-TTTTTTTT-LysNH ₂	2401.8	2397.3
R-PCNA1	H'-TTTTTTTT-LysNH ₂	2401.8	2402.0
S-PCNA2	H'-CTTCTTCCTT-LysNH ₂	2904.1	2907.5
S-PCNA3	H'-CATTGTCACACT-LysNH ₂	3524.1	3523.9
DNA1	5' GC AAA AAA AA CG 3' (complementarity to S-PCNA1/DNA4)		
DNA2	5' AAG AAA GAG A 3' (<i>ap</i> complementarity to S-PCNA2/DNA5)		
DNA3	5' TGG AGT GTG ACA ATG GTG 3' (<i>ap</i> complementarity to S-PCNA3/DNA6)		
DNA4	5' TTTTTTTT 3'		
DNA5	5' CTTCTTCCTT 3'		
DNA6	5' CACCATTGTCACACTCCA 3'		

3A.8 Biophysical studies of S-PCNA oligomers

Various biophysical techniques such as UV Job's plot, UV- T_m and gel electrophoretic assay were used to evaluate the properties of S-PCNA oligomers. The stoichiometry of the PCNA:DNA was first determined using Job's plot.¹⁹ The UV-melting studies were then carried out with all the synthesized oligomers using UV-Visible spectrophotometer. The complexation of PCNA1 to complementary DNA was confirmed by poly acrylamide gel electrophoresis (PAGE) using DNA as the control.

3A.8.1 Binding Stoichiometry: UV Job's plot

Ultraviolet absorption measurements are extremely useful to determine the stoichiometry of ODN:DNA/RNA complexes which was described by Job. In this method, the stoichiometry of the interacting strands could be obtained from the mixing curves, in which the absorbance at a given wavelength is plotted as a function of the mole fraction of

each strand which determines the stoichiometry of the complex information. Various stoichiometric mixtures of *S*-PCNA and DNA were made with relative molar ratios of strands 0:100, 10:90, 20:80, 30:70, 40:60, 50:50, 60:40, 70:30, 80:20, 90:10, 100:0, restricting the total strand concentration to 2 μM in 10mM sodium phosphate buffer of pH 7.2 (10 mM NaCl and 0.1 mM EDTA). A mixing curve was plotted, absorbance at fixed wavelength (λ_{max} 260 nm) against mole fraction of *S*-PCNA (Figure 9). The plot shows the change in absorbance value when the concentration of *S*-PCNA in the mixtures increased to 60 to 70%, which further increased beyond that proportion. Like in case of *R*-PCNA, this also supports the formation of polyprimidine *S*-PCNA:DNA complexes in 2:1 ratio and mixed purine and pyrimidine sequence are known to form only duplexes, hence they are taken in 1:1 ratio.

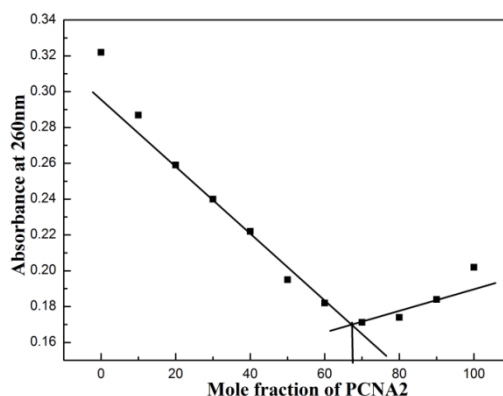


Figure 9 The Job's plot of UV- absorbance at 260 nm against Y-axis and the relative molar ratio of *S*-PCNA2/DNA2 (0→100) (10 mM sodium phosphate buffer, pH 7.2, 10 mM NaCl, 0.1 mM EDTA).

3A.8.2 UV- T_m - Results and discussion

Unlike DNA triplexes, which show two distinct transitions corresponding to triplex and duplex melting and later duplex to single strands melting, PCNA sequences exhibited single transition (triplex to single strand) similar to peptide nucleic acids. The hybridization studies of modified PCNAs with complementary DNA were done by temperature dependent UV-absorbance experiments. The stoichiometry PCNA:DNA complexation as established by UV mixing data at 260 nm (Job's plot) for

homopyrimidine was 2:1 ratio. PCNA hybridizes to complementary DNA/RNA sequences *via* specific base complementation to form duplexes for mixed sequences and triplexes for homopyrimidine/ homopurine sequences as in *R*-PCNA/PNA.²⁰

The UV-melting studies of *S*-PCNA with complementary DNA were measured on UV-Visible spectrophotometer. The UV melting studies revealed that the *S*-PCNA oligomers also, like *R*-PCNA, formed the very stable duplexes with the complementary DNA, Figure 10. The degree of stabilization was slightly higher in case of *S*-PCNA compared to *R*-PCNA oligomers.

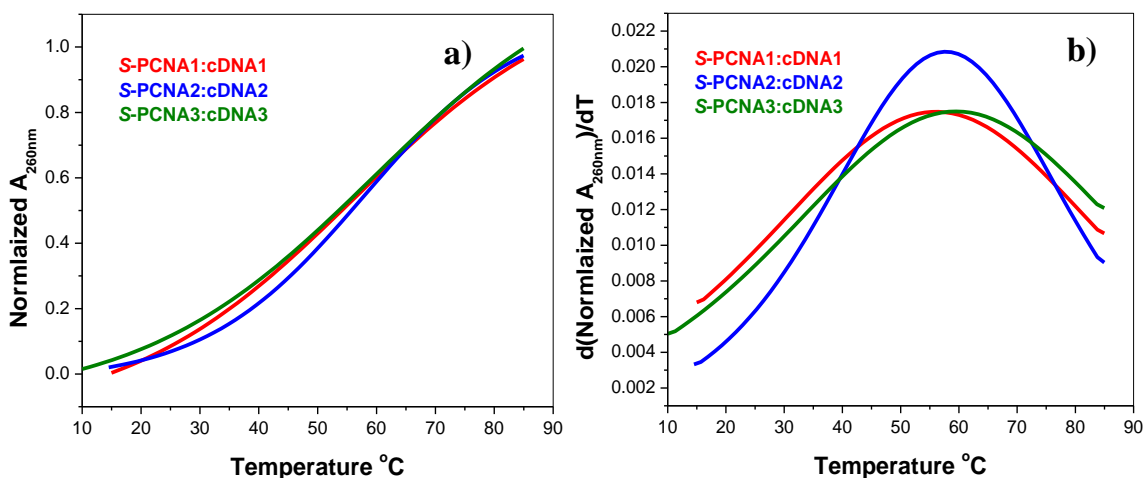


Figure 10 UV melting profiles (a) and their first order derivatives (b) of *S*-PCNA with cDNA.

The triplex UV-melting studies of octathyminyl sequences **S-PCNA1** and **R-PCNA1** (exhibiting the single transition) with the complementary **DNA1** (GC- clamps were provided at both 5' and 3' ends in order to retard the sliding) shown +34.7 and +28.6 $^{\circ}C$ stabilization (Figure 10), respectively, compared to the **DNA4-DNA1** control (Figure 11). **S-PCNA2** sequence (b_2a_2) is part of *bcr-abl*²¹ target gene and has a biological relevance. It is targeted against the gene causing chronic myeloid leukaemia (CML). The melting profiles observed for this poly pyrimidine sequence also follows the same pattern with **DNA2**. **S-PCNA2** binds to the **DNA2** also strongly with a stabilization of +26.4 $^{\circ}C$ at pH 7.2. After the poly pyrimidine sequences, the melting profiles of mixed sequence *S*-

Chapter 3

PCNA3 was also analyzed. **S-PCNA3**, a 12-mer, which is part of 18-mer micro-RNA sequence (three nucleotides were skipped at both 5' and 3' ends) also stabilizes the duplex (12mer:18mer) with the **DNA3** compared to the **DNA6:DNA3** (18mer:18mer, Figure 11).

Table 2 UV-melting studies of *S*-PCNA oligomers

Code	<i>S</i> -PCNA oligomers	UV- T_m °C* cDNA
S-PCNA1	H'-TTTTT TTT-LysNH ₂	56.2
R-PCNA1	H'-TTTTT TTT-LysNH ₂	50.1
S-PCNA2	H'-CTTCTTCCTT-LysNH ₂	57.6
S-PCNA3	H'-CATTGTCACACT-LysNH ₂	58.9 (12mer:18mer)
DNA4	5' TTTTTTTT 3'	21.5
DNA5	5' CTTCTTCCTT 3'	31.2
DNA6	5' CACCATTGTCACACTCCA 3'	53.9 (18mer:18mer)

*UV- T_m values were measured by annealing 1 μ M sequences with 1 μ M cDNA in sodium phosphate buffer (0.01M, pH 7.2) containing 10mM NaCl, 0.1mM EDTA and is an average of three independent experiments.

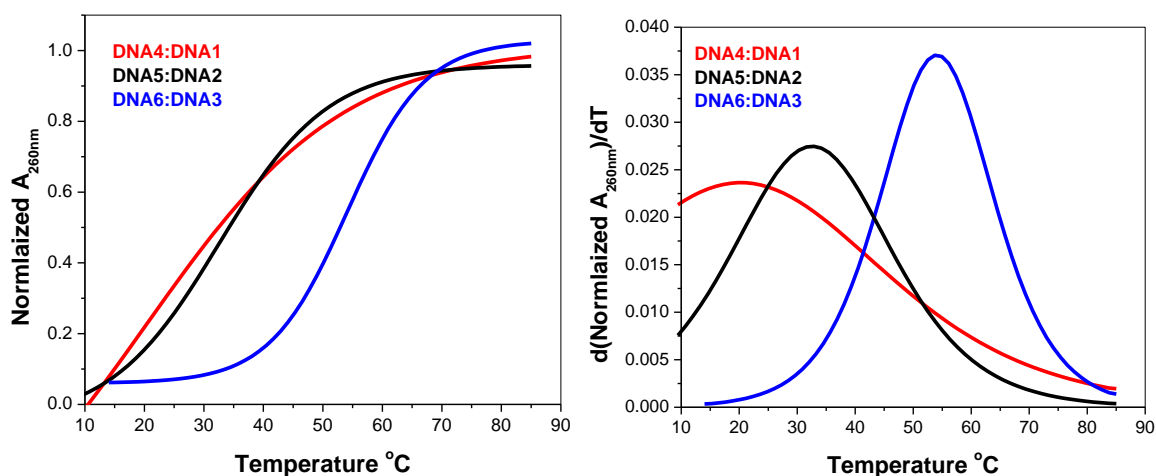


Figure 11 UV melting profiles and their first order derivatives of DNA with cDNA.

3A.8.3 Gel retardation of T₈ complexes

Electrophoretic gel experiment²² was used to establish the binding of **S-PCNA1** to the complementary **DNA1** (Figure 12). The complexation of **S-PCNA1:DNA1** was also planned to support with the complexation of **R-PCNA1:DNA1**, **DNA4:DNA1** as controls. The **PCNA1**, **R-PCNA1** and **DNA4** were mixed separately with **DNA1** in 2:1 ratio (T₈ strand, 0.7 mM and **DNA1**, 0.35 mM) in water. The samples were lyophilized to dryness and re-suspended in 2 μ L of 10mM sodium phosphate buffer containing 10mM NaCl and 0.1mM EDTA. The samples were annealed by heating to 85°C for 5 min followed by slow cooling to room temperature and refrigeration at 4°C overnight. To this, 2 μ L of 40% sucrose in TBE buffer pH 8.0 was added and the sample was loaded on the gel. Bromophenol blue (BPB) was used as the tracer dye separately in an adjacent well. Gel electrophoresis was performed on a 15% non-denaturing polyacrylamide gel (acrylamide:bis-acrylamide, 29:1) at constant power supply of 200V and 10 mA, until the BPB migrated to three-fourth of the gel length. During electrophoresis the temperature was maintained at ~5 °C. The bands were visualized through ethidium bromide staining of the gel.



Figure 12 Ethidium bromide staining of 15% non-denaturing PAGE analysis of complexes of **S-PCNA1**, **R-PCNA1** and **DNA4** with complementary **DNA 1**.

Chapter 3

Lane I corresponds to **R-PCNA1:DNA1** complex

Lane II corresponds to **S-PCNA1:DNA1** complex

Lane III single stranded **DNA1**

Lane IV corresponds to **DNA4:DNA1** complex

From the gel picture it was clearly understood the retardation of the bands in lane I and lane II indicating the formation of stable duplexes (neutral+charged) and lane IV indicating the formation of the weak DNA-DNA duplex (charged+charged). This experiment further confirmed the formation of complex of **S-PCNA1** with **DNA1**.

3A.9 Conclusion

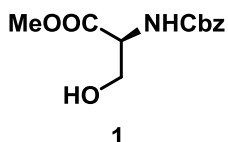
- ✓ We have successfully synthesized the enantiomeric **S-PCNA** oligomers from the same chiral source L-serine, as that of **R-PCNA**, by simply using the same set of chemical reactions but the only change being in the order of the chemical reactions.
- ✓ The concept of both the enantiomeric oligocarbamates from a single chiral source (without any inversion), L-serine, was ascertained through comparison of the specific rotation of the key intermediates.
- ✓ We performed the UV- T_m studies and observed that these **S-PCNA** oligomers also bind to DNA, like **R-PCNA** oligomers. Binding with RNA was not observed.
- ✓ The complex formation of these carbamate nucleic acids with cDNA were supported from polyacrylamide gel electrophoresis.

3A.10 Experimental Section

General: All the non-aqueous reactions were carried out under the inert atmosphere of Nitrogen/ Argon and the chemicals used were of laboratory or analytical grade. All solvents used were dried and distilled according to standard protocols. TLCs were carried out on pre-coated silica gel GF254 sheets (Merck 5554). All reactions were monitored by TLC and usual work-up implies sequential washing of the organic extract with water and brine followed by drying over anhydrous sodium sulfate and evaporation of the organic layer under vacuum. Column chromatographic separations were performed using silica gel 100-200 mesh (Merck) or 230- 400 mesh (Merck) and using the solvent systems EtOAc/Petroleum ether or MeOH/DCM. ^1H and ^{13}C NMR spectra were obtained using Bruker AC-200, AC-400 or AC-500 NMR spectrometers. The chemical shifts are reported in delta (δ) values and referred to internal standard TMS for ^1H . Optical rotations were measured on a JASCO DIP-181 polarimeter and mass spectra were recorded on a Finnigan-Matt mass spectrometer. Complementary DNA sequences were synthesized on Bioautomation Mer-Made 4 DNA synthesizer using standard β -cyanoethyl phosphoramidite chemistry. PCNA sequences were analyzed and purified by RP HPLC (Waters Delta 600e quaternary solvent delivery system and 2998 photodiode array detector and Empower2 chromatography software) on a C18 column using an increasing gradient of acetonitrile in water containing 0.1% TFA and characterized by MALDI-TOF mass spectrometry. (DNA sequences were also analyzed and purified under the same conditions with the increasing gradient of acetonitrile in 0.1N triethylammonium acetate of pH 7.0. The MALDI-TOF spectra were recorded on Voyager-De-STR (Applied Biosystems) MALDI-TOF or AB Sciex TOF/TOFTM series explorer TM 72085 instrument; the matrix used for PCNA samples was α -cyanohydroxycinnamic acid and THAP (2, 4, 6-trihydroxyacetophenone) for DNA samples. Thermal denaturation experiments were performed on Varian Cary-300 UV-Vis spectrophotometer fitted with a peltier-controlled temperature programmer and a water circulator, at the temperature ramping rate 0.5 $^\circ\text{C}/\text{min}$ and the absorbance was recorded at 260 nm for every 0.5 $^\circ\text{C}$ rise in temperature.

Methyl ((benzyloxy)carbonyl)-L-serinate, 1:

To a solution of L-serine (476 mmol, 50 g) and NaHCO₃ (1190 mmol, 100 g) in water (1 L), Cbz-Cl (571 mmol, 81.7 mL) was added at 0 °C and the reaction mixture was kept

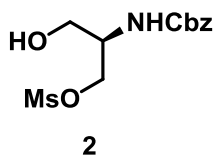


vigorously stirring at room temperature overnight. The reaction mixture was neutralized with 2N HCl and the pH was maintained at ~3. The product started precipitating out as a white solid. The reaction mixture was filtered and the white solid was desiccated to yield Cbz-

protected L-serine (80 g at 70% yield). To Cbz-protected L-serine (62.7 mmol, 15g) in MeOH (120 mL) was added NEt₃ (156.7 mmol, 21.8 mL) followed by the drop wise addition of SOCl₂ (75.2 mmol, 5.48 mL) at 0 °C. The reaction mixture was brought to room temperature and refluxed for 6h. MeOH was removed *in vacuo* and the residue was diluted with EtOAc; water wash, aq. NaHCO₃ wash and brine wash were given. The organic layer was dried over anhydrous Na₂SO₄ and concentrated *in vacuo*. The crude compound was column purified. Compound **1** eluted in 45% EtOAc in petroleum ether as a colorless liquid in 90% yield. ¹H NMR (200 MHz, CDCl₃): δ 2.39 (t, 1H, J=5.9 Hz), 3.78 (s, 3H), 3.95 (m, 2H), 4.47 (m, 1H), 5.12 (s, 2H), 5.74 (d, 1H, J=6.48 Hz), 7.35 (m, 5H); ¹³C NMR (50 MHz, CDCl₃): δ 52.1, 55.8, 62.1, 66.6, 127.5, 128.0, 135.7, 156.1, 171.0; MS (EI): Mass calculated for C₁₂H₁₅NO₅Na (M+Na) 276.09, found 276.05.

(R)-2-(((benzyloxy)carbonyl)amino)-3-hydroxypropyl methanesulfonate, 2:

To a solution of NH-Cbz-L-serine methyl ester **1** (19.8 mmol, 5 g) in dry pyridine (20 mL), Ms-Cl (39.5 mmol, 3 mL) was added at 0 °C. The reaction was complete within 15 minutes.



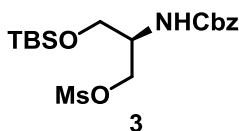
The reaction mixture was diluted with CH₂Cl₂ (300 mL) and was washed with 5% aq. KHSO₄ and brine solution, the organic layer was dried over anhydrous Na₂SO₄. CH₂Cl₂ was evaporated *in vacuo* at rt (to avoid the thermal decomposition of mesylate-ester to unwanted α, β-

unsaturated methyl ester) and the crude mesylate was directly subjected to NaBH₄ reduction. The residue was dissolved in MeOH (150 mL) and NaBH₄ (12.4 g, 32.64 mmol) was added in portions over a period of 30 min at 0 °C. The reaction mixture was stirred for

a period of 3 h at rt and finally quenched with saturated NH_4Cl solution till pH was neutral. Methanol was removed under reduced pressure and the residue was extracted into EtOAc (3 x 150 mL). The organic layer was dried over Na_2SO_4 and solvent was removed under reduced pressure. The residue was purified by column chromatography (eluted in 40% EtOAc in petroleum ether) affording compound **2** in 75% yield over two steps. $^1\text{H NMR}$ (200 MHz, CDCl_3): δ 2.68 (bs, 2H), 2.99(s, 3H), 3.64-3.76 (m, 2H), 3.94-4.03 (m, 1H), 4.3 (m, 2H), 5.09 (s, 2H), 5.5 (s, 1H), 7.34 (s, 5H); $^1\text{H NMR}$ (200 MHz, $\text{CDCl}_3 + \text{D}_2\text{O}$): δ 2.98 (s, 3H), 3.6-3.74 (m, 2H), 3.97 (pentet, $J=4.87, 9.72$ Hz, 1H), 4.28 (m, 2H), 4.90 (bs, 2H), 5.08 (s, 2H), 7.33 (s, 5H).

(R)-2-(((benzyloxy)carbonyl)amino)-3-((tert-butyldimethylsilyl)oxy)propylmethane sulfonate, 3:

Alcohol **2** (29.7 mmol, 9.0 g) was dissolved in dry CH_2Cl_2 (40 mL) and TBS-Cl (35.6 mmol, 5.4 g), imidazole (74.2 mmol, 5.0 g) were added and the reaction mixture was kept for stirring at rt for 5 h. The reaction mixture was diluted with the CH_2Cl_2 (300 mL) and was washed with water and brine solutions. The CH_2Cl_2 layer was dried over Na_2SO_4

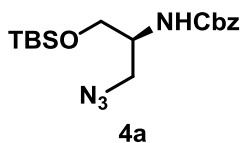


and solvent was removed *in vacuo*. Crude compound was column purified (eluted in 15% EtOAc in petroleum ether) to furnish **3** as a white solid in 96% yield. $^1\text{H NMR}$ (200 MHz, CDCl_3): δ 0.07 (s, 6H), 0.88 (s, 9H), 3.0 (s, 3H), 3.6-3.8 (m, 2H), 4.01 (m, 1H), 4.27-4.29 (m, 2H), 5.11 (s, 2H), 7.36 (m, 5H); $^{13}\text{C NMR}$ (50 MHz, CDCl_3): δ -5.6, 18.1, 25.7, 37.1, 51.1, 60.9, 67.0, 67.3, 128.1, 128.5, 136.1, 155.7; **MS (EI)**: Mass calculated for $\text{C}_{18}\text{H}_{31}\text{NO}_6\text{SSiNa}$ ($M+\text{Na}$) 440.16, found 440.47.

and solvent was removed *in vacuo*. Crude compound was column purified (eluted in 15% EtOAc in petroleum ether) to furnish **3** as a white solid in 96% yield. $^1\text{H NMR}$ (200 MHz, CDCl_3): δ 0.07 (s, 6H), 0.88 (s, 9H), 3.0 (s, 3H), 3.6-3.8 (m, 2H), 4.01 (m, 1H), 4.27-4.29 (m, 2H), 5.11 (s, 2H), 7.36 (m, 5H); $^{13}\text{C NMR}$ (50 MHz, CDCl_3): δ -5.6, 18.1, 25.7, 37.1, 51.1, 60.9, 67.0, 67.3, 128.1, 128.5, 136.1, 155.7; **MS (EI)**: Mass calculated for $\text{C}_{18}\text{H}_{31}\text{NO}_6\text{SSiNa}$ ($M+\text{Na}$) 440.16, found 440.47.

Benzyl (S)-(1-azido-3-((tert-butyldimethylsilyl)oxy)propan-2-yl)carbamate, 4a:

A stirred mixture of mesylate **3** (8.3 mmol, 5.0 g) and NaN_3 (99.6 mmol, 6.4 g) in dry DMF (20 mL) was heated at 60 °C for a period of 5 h. DMF was removed *in vacuo* at 40 °C and the residue was extracted into EtOAc (3x 100 mL).

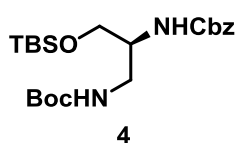


The combined organic layer was washed with water and brine and

dried over Na_2SO_4 . The residue was further purified by column chromatography (eluted in 7% EtOAc in petroleum ether) to afford azide **5** as a colorless oil in 76% yield. **IR** (CHCl_3 , v/cm^{-1}): 3334, 2954, 2929, 2102, 1724, 1255; **^1H NMR** (200 MHz, CDCl_3): δ 0.04 (s, 6H), 0.87 (s, 9H), 3.36-3.52 (m, 2H), 3.54-3.72 (m, 2H), 3.79-3.83 (m, 1H), 5.08 (s, 2H), 5.17 (d, $J = 8.38$ Hz), 7.31 (m, 5H); **^{13}C NMR** (50 MHz, CDCl_3): δ -5.7, 18.0, 25.6, 50.8, 51.5, 61.6, 66.7, 127.92, 127.97, 128.3, 136.1, 155.6.

Benzyl tert-butyl (3-((tert-butyldimethylsilyl)oxy)propane-1,2-diyl)(S)-dicarbamate, 4:

To a solution of azide **4a** (6 g, 16.4 mmol) in EtOAc (25 mL) placed in hydrogenation flask

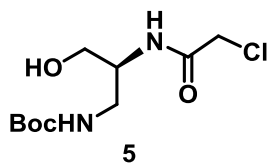


was added di-*tert*-butyl dicarbonate (4.9 mL, 21.4 mmol) and Raney-nickel (10% w/W). The mixture was subjected to hydrogenation in Parr apparatus at room temperature (40 psi) for 3 h. The slurry was filtered through celite and the filtrate was evaporated under reduced

pressure. The residue was purified by column chromatography (eluted in 7% EtOAc in petroleum ether) to afford compound **4** as colorless oil in 84% yield over two steps. **^1H NMR** (200 MHz, CDCl_3): δ 0.04 (s, 6H), 0.88 (s, 9H), 1.42 (s, 9H), 3.27-3.36 (m, 2H), 3.62-3.77 (m, 3H), 5.1 (s, 2H), 5.34-5.38 (m, 1H), 7.34 (m, 5H); **^{13}C NMR** (50 MHz, CDCl_3): δ -5.5, 18.2, 25.8, 28.3, 42.5, 52.5, 63.6, 66.7, 79.4, 128.0, 128.5, 136.5, 156.6; **MS (EI)**: Mass calculated for $\text{C}_{22}\text{H}_{38}\text{N}_2\text{O}_5\text{SiNa}$ ($\text{M}+\text{Na}$) 460.25, found 460.86.

tert-butyl (S)-(2-(2-chloroacetamido)-3-hydroxypropyl)carbamate, 5:

To a solution of compound **4** (11.4 mmol, 5 g) in methanol (15 mL) placed in hydrogenation flask was added 10% Pd/C (0.5 g, 10% w/W). The



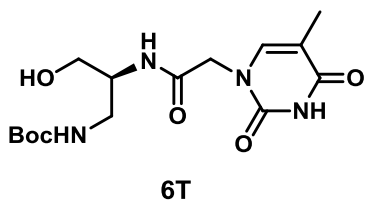
mixture was hydrogenated in Parr apparatus (60 psi) at room temperature for 6 h. The slurry was filtered through celite and the filtrate was evaporated under reduced pressure. The crude amine

was subjected to chloroacetylation with chloroacetyl chloride. To a stirred solution of the crude amine **4** (11.9 mmol, 3.62 g) in 1,4-dioxane (10 mL) was added K_2CO_3 (84 mmol, 11.6 g) in water (20 mL) and followed by 1,4-dioxane (10 mL). The mixture was cooled to 0°C , and chloroacetyl chloride (60.0 mmol, 4.77 mL) was added portion wise. The reaction

mixture was kept stirring at rt for 6h. Dioxane was removed under reduced pressure and the residue was extracted with EtOAc (2 x100 mL). The organic layer was dried over Na₂SO₄ and the solvent was evaporated under reduced pressure. The residue was purified by column chromatography (eluted in 40% EtOAc in petroleum ether) affording the chloro derivative **5** as colorless oil in 80% yield. ¹H NMR (200 MHz, CDCl₃): δ 1.44 (s, 9H), 3.15-3.57 (m, 4H), 3.7-3.78 (m, 1H), 4.03 (s, 2H), 5.14 (t, 1H, J= 6.3, 12.8 Hz); ¹³C NMR (50 MHz, CDCl₃): δ 28.2, 39.7, 42.4, 52.0, 60.9, 80.4, 157.7, 166.5; MS (EI): Mass calculated for C₁₀H₁₉ClN₂O₄Na (M+Na) 289.10, found 289.41.

tert-butyl (S)-(3-hydroxy-2-(2-(N1-thyminy)acetamido)propyl)carbamate, 6T:

To a solution of compound **5** (6.0 mmol, 1.6 g) in dry DMF (8 mL), thymine (6.0 mmol, 0.76 g) and anhydrous K₂CO₃ (6.66 mmol, 0.92 g) were added. The reaction mixture was stirred at room temperature overnight under nitrogen. The solvent was removed under reduced pressure and the residue was extracted in EtOAc (100 mL x 2) and dried over Na₂SO₄. The solvent was

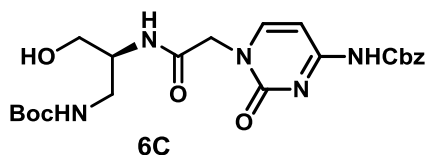


evaporated and the crude compound was purified by column chromatography (eluted in 80%EtOAc in petroleum ether) to furnish thymine monomer **6T** as a white solid in 91% yield. IR (CHCl₃, ν/cm⁻¹): 3427, 3106, 2954, 1712, 1510, 1215, 757; ¹H NMR (200 MHz, CDCl₃): δ 1.37 (s, 9H), 1.75 (s, 3H), 2.91-3.11 (m, 2H), 3.74 (m, 1H), 4.27 (s, 2H), 4.67 (t, J=5.69 Hz, 1H), 6.72-6.78 (m, 1H), 7.39 (m, 1H), 7.91 (d, J=8.09 Hz, 1H) ¹³C NMR (50 MHz, CDCl₃): δ 11.9, 28.2, 49.2, 51.4, 54.9, 60.8, 77.9, 107.9, 142.3, 151.0, 156.0, 164.5, 166.8; MS (EI): Mass calculated for C₁₅H₂₄N₄O₆Na (M+Na) 379.16, found 379.28.

tert-butyl (S)-(3-hydroxy-2-(2-(N1-cytosinyl^{Cbz})acetamido)propyl)carbamate, 6C:

A mixture of compound **5** (9.0 mmol, 2.4 g), N-Cbz-cytosine (9.0 mmol, 2.2 g) and anhydrous K₂CO₃ (10.9 mmol, 1.5 g) in dry DMF (10 mL) was stirred at 50 °C for 12 h under nitrogen. The solvent was removed under reduced pressure and the residue was extracted into EtOAc (200 mL x 2) and dried over Na₂SO₄. The solvent was evaporated and

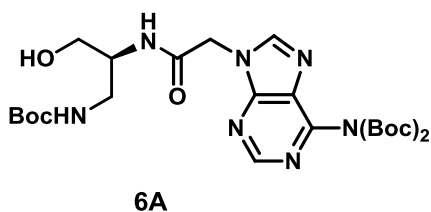
the crude compound was purified by column chromatography (eluted in 75%EtOAc in petroleum ether) to afford **6C** as a white foam in 80% yield.



IR (CHCl_3 , v/cm^{-1}): 3432, 3110, 2958, 17102, 1510, 1215, 757; **^1H NMR** (200 MHz, CDCl_3): δ 1.33 (s, 9H), 3.21 (m, 2H), 3.51 (m, 2H), 3.88 (m, 1H), 4.45 (s, 2H), 5.13 (s, 2H), 5.73-5.76 (m, 1H), 7.30 (m, 5H), 7.58 (d, 1H, $J=7.29$ Hz), 7.77-7.81 (m, 1H); **^{13}C NMR** (50 MHz, DMSO-d_6): δ 28.2, 51.5, 60.8, 66.5, 77.9, 93.8, 128.0, 128.2, 128.5, 136.0, 151.0, 153.2, 155.2, 156.0, 163.1, 166.6; **MS (EI)**: Mass calculated for $\text{C}_{22}\text{H}_{29}\text{N}_5\text{O}_7\text{Na}$ ($\text{M}+\text{Na}$) 498.20, found 498.63.

tert-butyl (S)-(3-hydroxy-2-(2-(N9-adeninyl^{(Boc)₂})acetamido)propyl) carbamate, 6A:

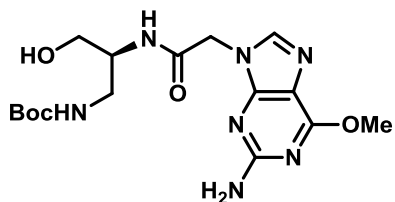
To a solution of compound **5** (8.6 mmol, 2.3 g) in dry DMF (10 mL), $(\text{Boc})_2$ -adenine (8.6 mmol, 2.9 g) and anhydrous K_2CO_3 (10.3 mmol, 1.4 g) were added. The reaction mixture was stirred at room temperature overnight under nitrogen atmosphere. The solvent was removed under reduced pressure and the residue was extracted into EtOAc (200 mL x 2) and



dried over Na_2SO_4 . The solvent was evaporated and the crude compound was purified by column chromatography (eluted in 65%EtOAc in petroleum ether) to furnish thymine monomer **6A** as white solid in 75% yield. **^1H NMR** (200 MHz, CDCl_3): δ 1.45 (s, 27H), 3.16-3.41 (m, 2H), 3.50-3.75 (m, 2H), 3.88 (m, 1H), 4.92 (s, 2H), 5.03-5.10 (m, 1H), 5.30 (s, 2H), 7.10 (d, $J = 8.05$ Hz, 1H), 8.21 (s, 1H), 8.85 (s, 1H); **MS (EI)**: Mass calculated for $\text{C}_{25}\text{H}_{39}\text{N}_7\text{O}_8\text{Na}$ ($\text{M}+\text{Na}$) 588.28, found 588.28.

tert-butyl (S)-(3-hydroxy-2-(2-(9-(6-O-methyl)guaninyl)acetamido)propyl)carbamate, 6G:

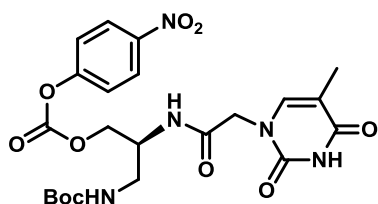
A mixture of compound **5** (4.3 mmol, 1.5 g), 2-amino-6-chloropurine (4.7 mmol, 1.0 g) and anhydrous K_2CO_3 (5.1 mmol, 0.97 g) in dry DMF (10 mL) was stirred at room temperature for 12 h under nitrogen. The solvent was removed under reduced pressure and the residue

**6G**

was extracted to EtOAc (200 mL x 2) and dried over Na₂SO₄. The solvent was evaporated and the crude compound was purified by column chromatography (eluted in 80% EtOAc in petroleum ether) to afford crude **6G^{Cl}** as a white solid and was directly subjected to basic hydrolysis to yield the guanine- monomer. The crude compound was dissolved in MeOH (10 mL) and 0.75 N NaOH (10 mL) was added. The reaction mixture was allowed to stir for 3.5 h. The reaction mixture was directly extracted into EtOAc (200 mL x 2), washed with water, brine and dried over Na₂SO₄. The solvent was evaporated and the crude compound was purified by column chromatography (eluted in 80% EtOAc in petroleum ether) to afford colorless flakes of compound **6G** in 69% yield over two steps. **IR** (CHCl₃, v/cm⁻¹): 3434, 3116, 2964, 1712, 1504, 1215, 757; **¹H NMR** (200 MHz, CDCl₃+ DMSO-d₆): δ 1.37 (s, 9H), 2.91-3.19 (m, 3H), 3.7-3.79 (m, 2H), 4.69 (s, 2H), 6.36 (s, 2H), 6.77 (t, J=5.65 Hz, 1H), 8.07 (d, J=7.98 Hz, 1H); **¹³C NMR** (50 MHz, DMSO-d₆): δ 28.2, 44.8, 51.5, 53.2, 60.8, 77.9, 113.3, 140.7, 154.4, 156.0, 159.8, 160.6, 166.4; **MS (EI)**: Mass calculated for C₁₆H₂₅N₇O₅Na (M+Na) 418.19, found 418.26.

***tert*-butyl (S)-(2-(2-(N1-thyminy)acetamido)-3-(oxycarbonyl(4-nitrophenoxy))propyl) carbamate, 7T:**

Compound **6T** (4.2 mmol, 1.5 g) was taken in dry DCM (15 mL) and cooled to 0 °C in an ice bath. Dry pyridine (8.42 mmol, 0.67 mL) was added and allowed to stir for 10 min. *p*-nitrophenylchloroformate (6.3 mmol, 1.2 g) dissolved in DCM (5 mL) was added and the reaction allowed stirring at room temperature for 2 h. After completion of reaction, the reaction mixture was extracted into DCM (100 mL x 2), washed with water, brine and dried over Na₂SO₄. The solvent was evaporated and the crude compound was purified by column chromatography (eluted in 75% EtOAc in petroleum ether) to afford white activated thymine- monomer **7T** in 70% yield along with the recovery of **6T** in 12% yield. **¹H NMR** (200 MHz, CDCl₃): δ 1.41 (s, 9H), 1.87 (s, 3H), 3.37 (m, 2H), 4.32 (m,

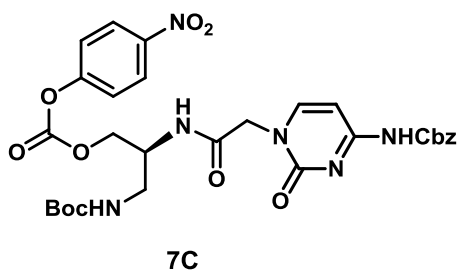
**7T**

Chapter 3

4H), 5.23 (bs, 1H), 7.06 (s, 1H), 7.35-7.40 (d, $J = 9.13$ Hz, 2H), 8.24-8.28 (d, $J = 9.08$ Hz, 2H); **MS (EI)**: Mass calculated for $C_{22}H_{27}N_5O_{10}Na$ ($M+Na$) 544.17, found 544.68.

tert-butyl (S)-(2-(2-(N1-cytosinyl^{Cbz})acetamido)-3-(oxycarbonyl(4-nitrophenoxy)) propyl)carbamate, 7C:

Compound **6C** (3.1 mmol, 1.5 g) was taken in dry DCM (15 mL) and cooled to 0 °C in an ice bath. Dry pyridine (6.3 mmol, 0.50 mL) was added and allowed to stir for 10 min. *p*-

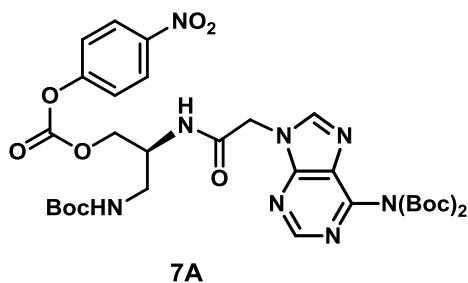


p-nitrophenylchloroformate (4.7 mmol, 0.95 g) dissolved in DCM (5 mL) was added and the reaction allowed to stir at room temperature for 1.5 h. After completion of reaction, the reaction mixture was extracted into DCM (100 mL x 2), washed with water, brine and dried over Na_2SO_4 . The solvent was

evaporated and the crude compound was purified by column chromatography (eluted in 70%EtOAc in petroleum ether) to afford white activated cytosine- monomer **7C** in 72% yield. 1H NMR (200 MHz, $CDCl_3$): δ 1.46 (s, 9H), 3.25-3.49 (m, 2H), 4.20-4.42 (m, 4H), 4.50-4.67 (m, 1H), 5.19 (s, 2H), 7.33-7.37 (bs, 7H), 7.64-7.84 (m, 2H), 8.20-8.25 (d, $J = 9.09$ Hz, 2H); **MS (EI)**: Mass calculated for $C_{29}H_{32}N_6O_{11}Na$ ($M+Na$) 663.21, found 663.09.

tert-butyl (S)-(2-(2-(N9-adeninyl^{(Boc)₂})acetamido)-3-(oxycarbonyl(4-nitrophenoxy)) propyl) carbamate, 7A:

Compound **6A** (3.0 mmol, 1.7 g) was taken in dry DCM (15 mL) and cooled to 0 °C in an



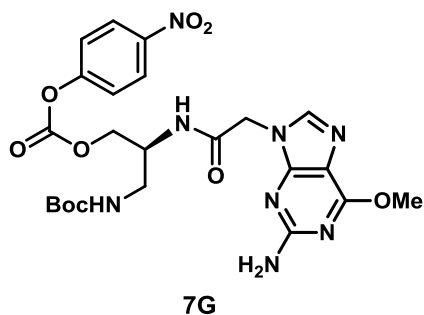
ice bath. Dry pyridine (6.1 mmol, 0.49 mL) was added and allowed to stir for 10 min. *p*-nitrophenylchloroformate (4.6 mmol, 0.93 g) dissolved in DCM (5 mL) was added and the reaction allowed to stir at room temperature for 100 min. After completion of reaction, the reaction

mixture was extracted into DCM (100 mL x 2), washed with water, brine and dried over

Na₂SO₄. The solvent was evaporated and the crude compound was purified by column chromatography (eluted in 60%EtOAc in petroleum ether) to afford white activated adenine- monomer **7A** in 86% yield. ¹H NMR (200 MHz, CDCl₃): δ 1.44 (s, 27H), 2.13 (m, 1H), 3.39 (m, 2H), 4.32 (m, 3H), 4.95 (s, 2H), 5.22 (t, J = 5.64 Hz, 1H), 7.33-7.38 (m, 2H), 7.57(m, 1H), 8.21 (s, 1H), 8.23-8.28 (m, 2H), 8.81 (s, 1H); ¹³C NMR (50 MHz, CDCl₃): δ 27.7, 28.2, 40.7, 46.0, 50.1, 67.4, 80.4, 84.0, 115.6, 121.7, 125.2, 126.0, 128.2, 145.4, 150.4, 152.2, 153.4, 155.2, 157.2, 165.9; MS (EI): Mass calculated for C₃₂H₄₃N₈O₁₂ (M+H) 731.29, found 731.65.

tert-butyl (S)-(2-(2-(9-(6-O-methyl)guaninyl)acetamido)-3-(oxycarbonyl(4-nitrophenoxy))propyl)carbamate, 7G:

Compound **6G** (4.2 mmol, 1.6 g) was taken in dry DCM (18 mL) and cooled to 0 °C in an

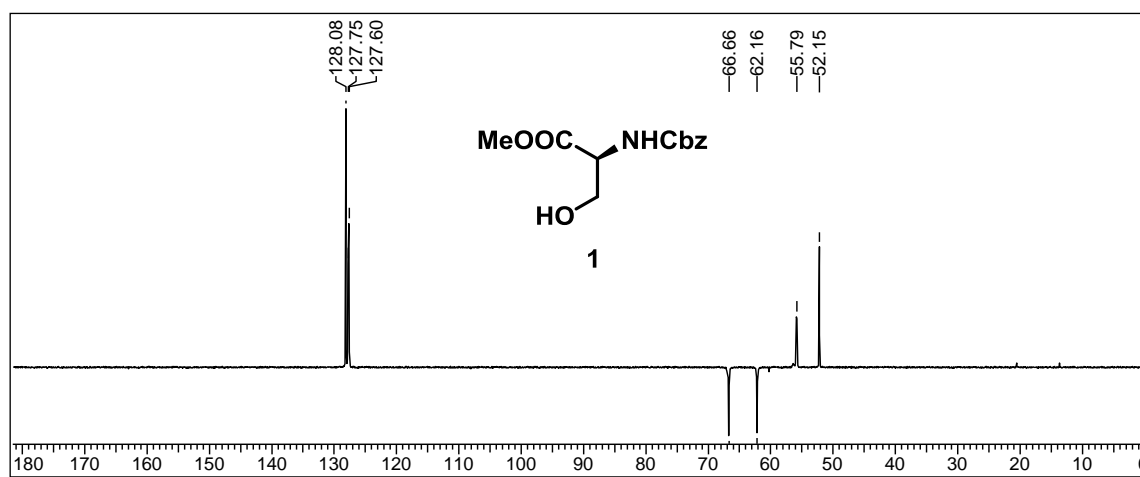
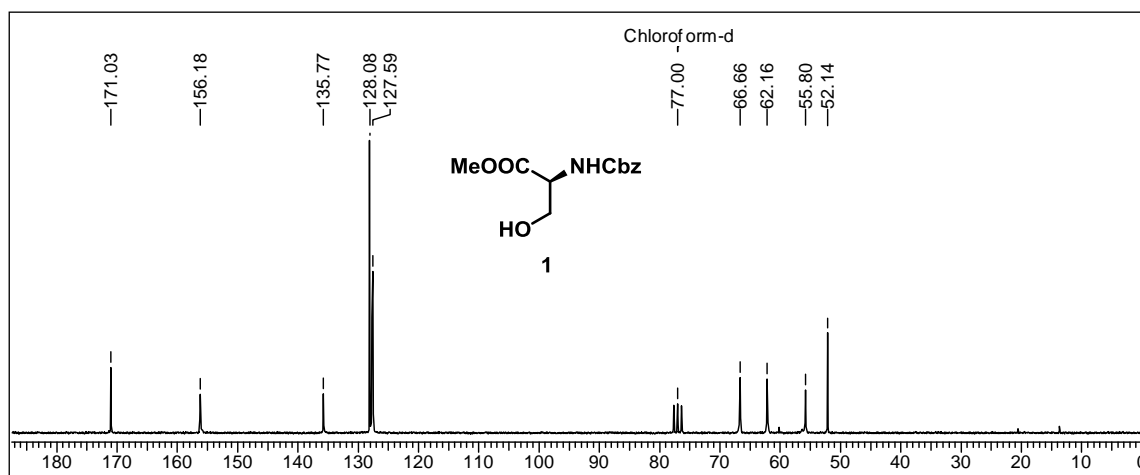
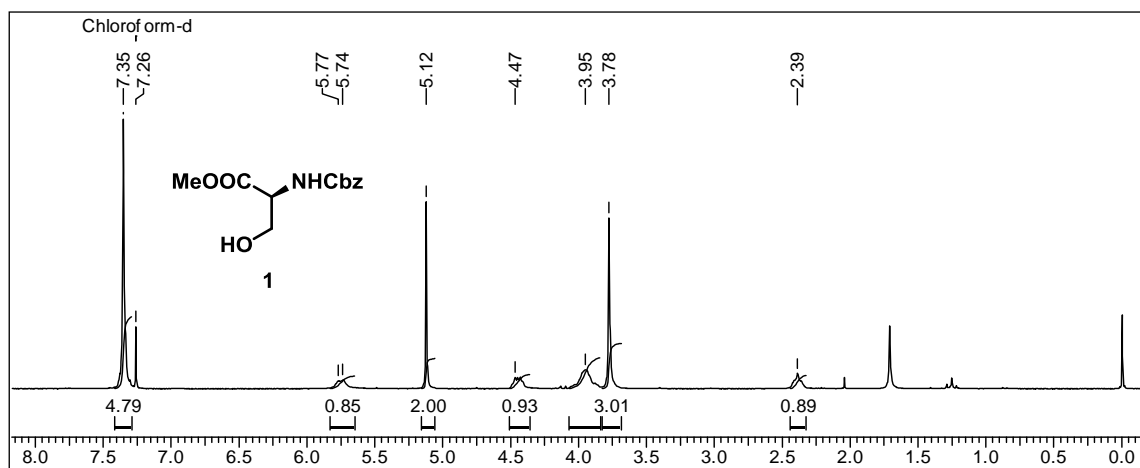


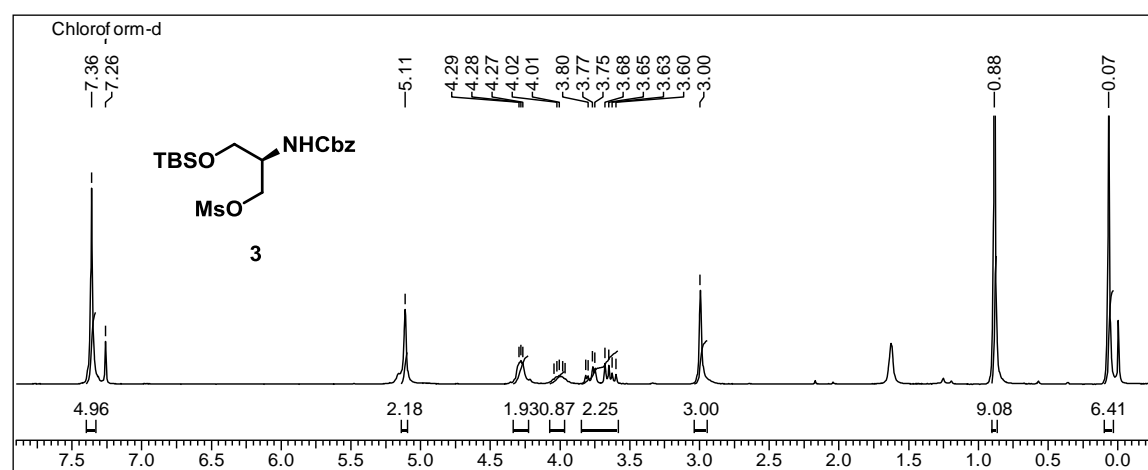
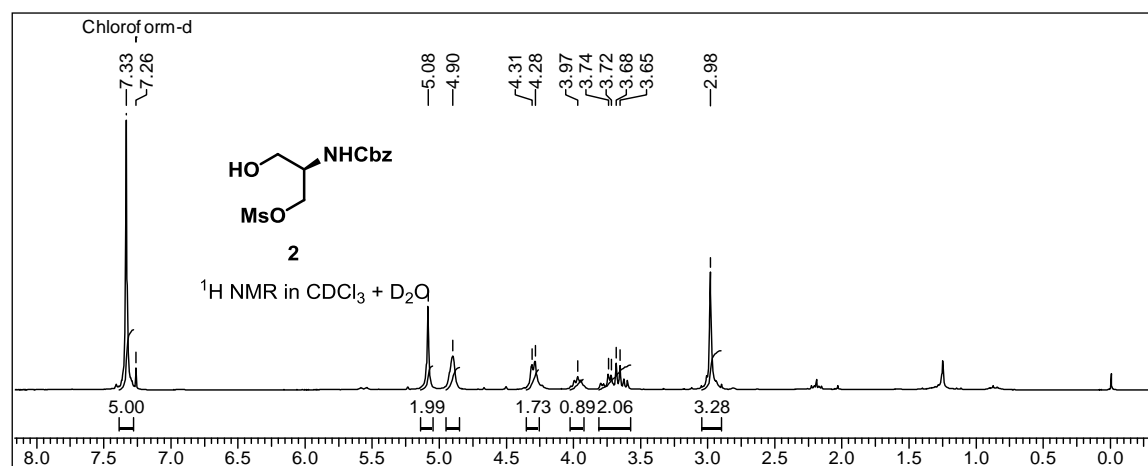
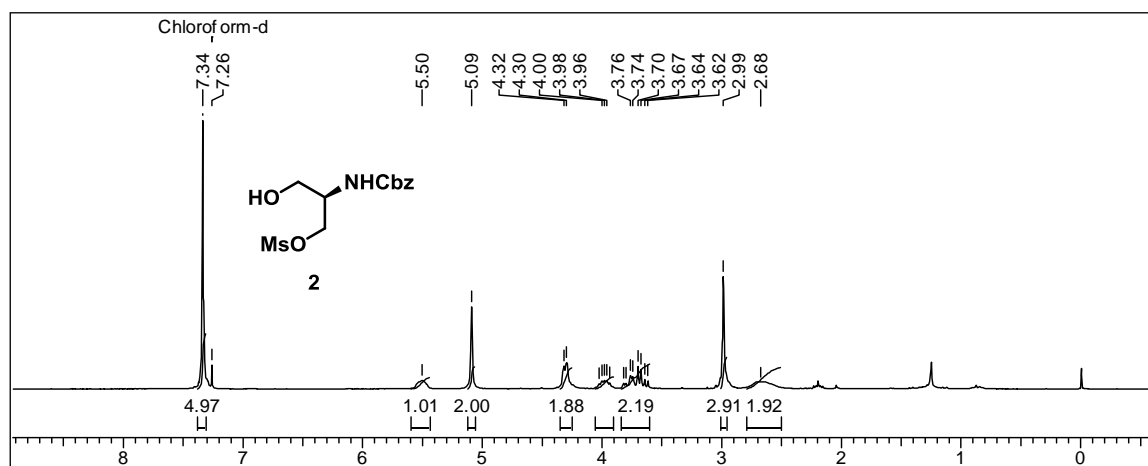
ice bath. Dry pyridine (8.4 mmol, 0.68 mL) was added and allowed to stir for 10 min. *p*-nitrophenylchloroformate (6.3 mmol, 1.2 g) dissolved in DCM (5 mL) was added and the reaction allowed to stir at room temperature for 2h. After completion of reaction, the reaction mixture was extracted into DCM (100 mL x 2), washed with water, brine and dried over

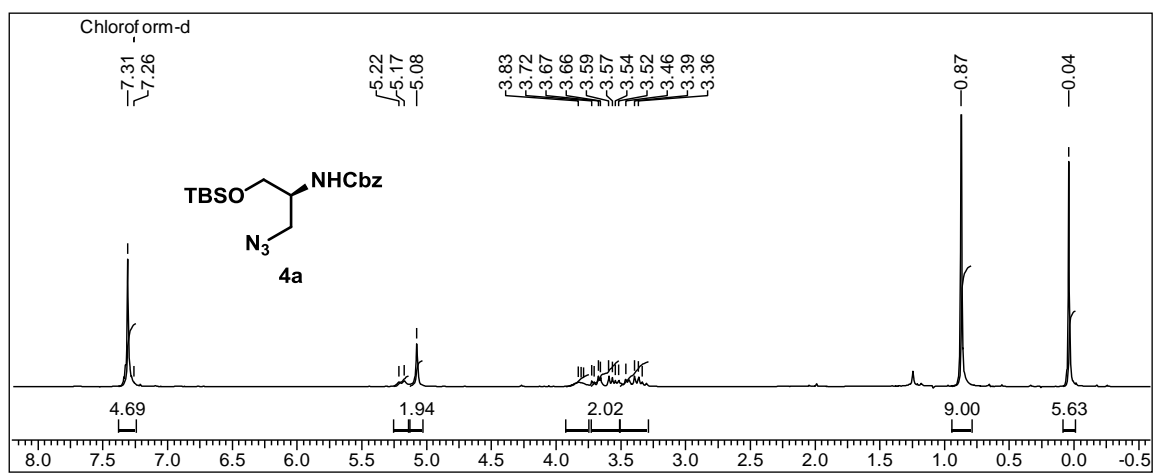
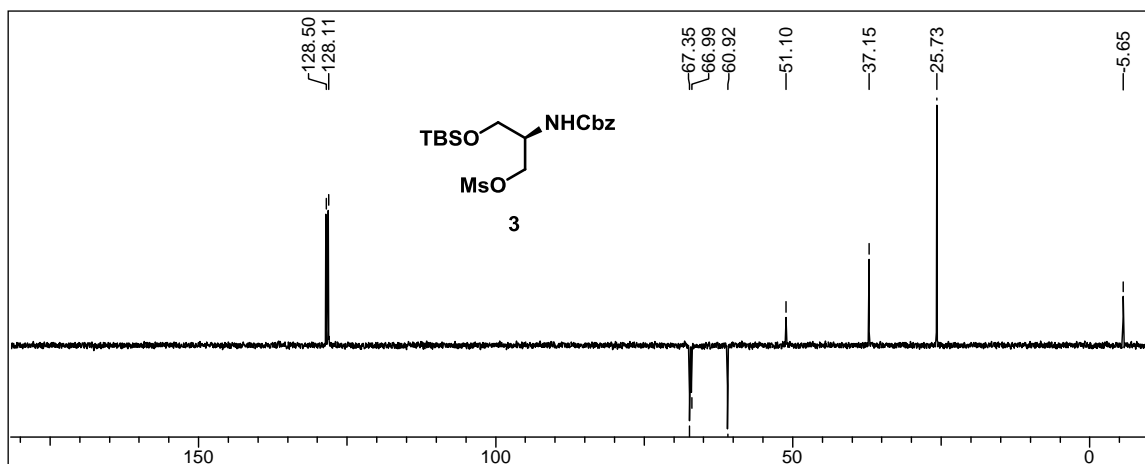
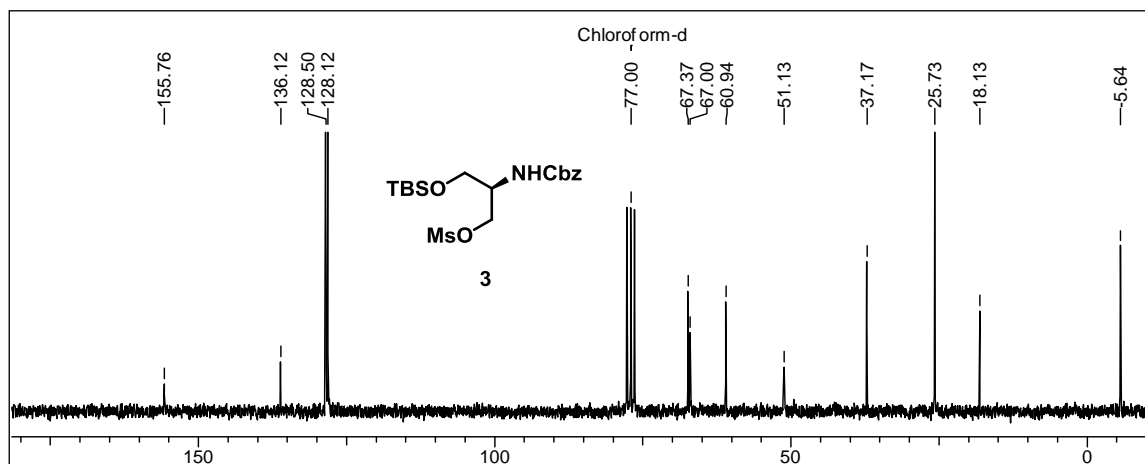
Na₂SO₄. The solvent was evaporated and the crude compound was purified by column chromatography (eluted in 75%EtOAc in petroleum ether) to afford white activated guanine- monomer **7G** in 66% yield along with the recovery of 17% **6G**. ¹H NMR (200 MHz, CDCl₃): δ_H 1.42 (s, 9H), 3.28-3.46 (m, 2H), 4.03 (s, 3H), 4.31 (bs, 3H), 4.71 (s, 2H), 7.33-7.38 (d, J = 9.14 Hz, 2H), 7.63 (m, 1H), 7.85 (m, 1H), 8.26-8.3 (d, J = 9.16 Hz, 2H); MS (EI): Mass calculated for C₂₃H₂₈N₈O₉ (M+H) 561.19, found 561.15.

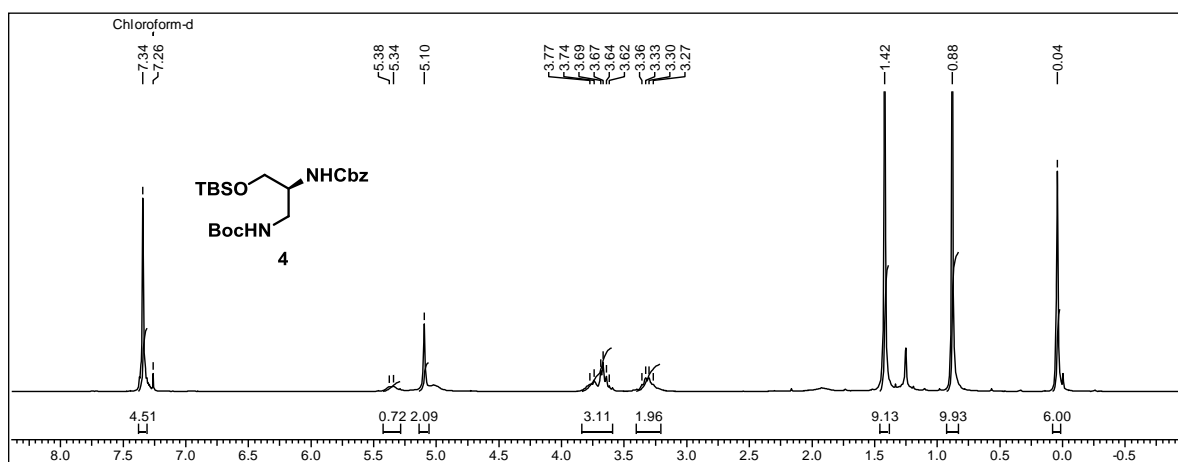
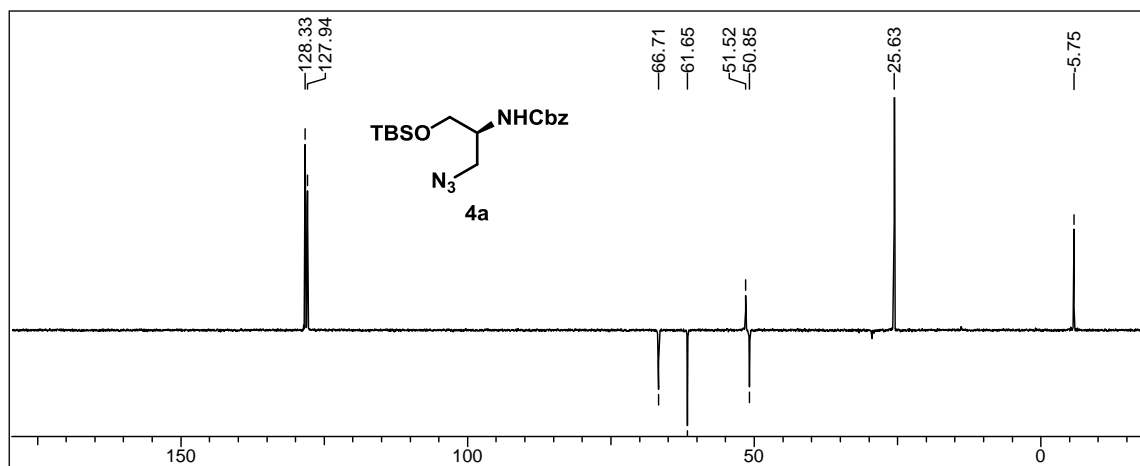
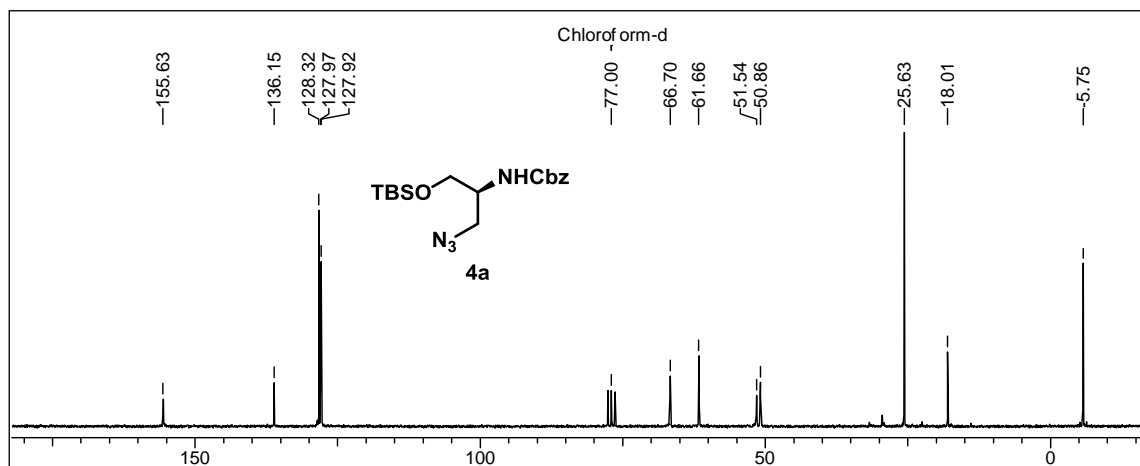
3A.11 Appendix

Compounds - Spectral data	Page No.
1- ^1H , ^{13}C NMR & DEPT	203
2- ^1H NMR & D_2O exchange, 3- ^1H NMR	204
3- ^{13}C NMR & DEPT, 4a- ^1H NMR	205
4a- ^{13}C NMR & DEPT, 4- ^1H NMR	206
4- ^{13}C NMR & DEPT, 5- ^1H NMR	207
5- ^{13}C NMR & DEPT, 6T- ^1H NMR	208
6T- ^{13}C NMR & DEPT, 6C- ^1H NMR	209
6C- ^{13}C NMR & DEPT, 6A- ^1H NMR	210
6G- ^1H , ^{13}C NMR & DEPT	211
7T, 7C & 7G- ^1H NMR	212
7A- ^1H , ^{13}C NMR & DEPT	213
1, 3- LC MS	214
4, 5- LC MS	215
6T, 6C- LC MS	216
6A, 7T- LC MS	217
7C, 7A- LC MS	218
7G- LC MS, PCNA1- MALDI-TOF spectra	219
PCNA2 & PCNA3- MALDI-TOF spectra	220

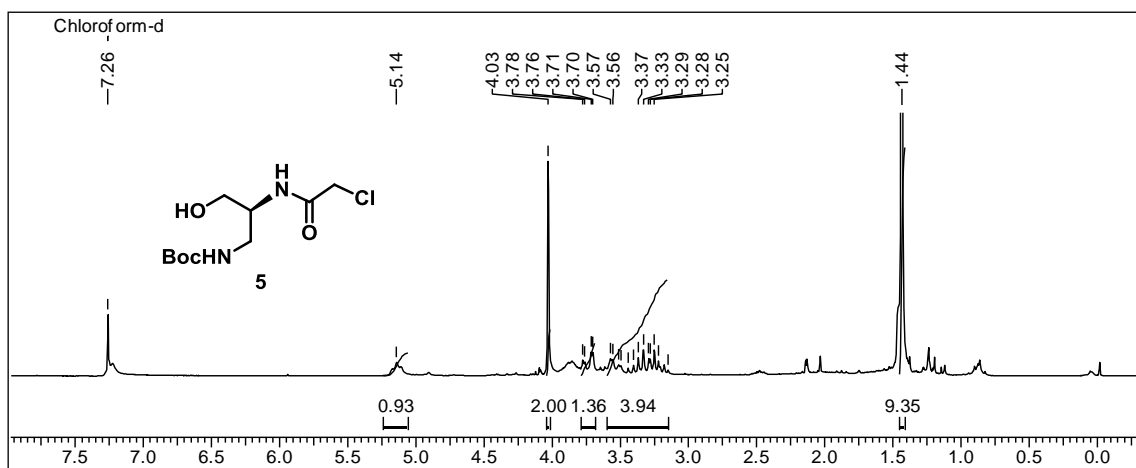
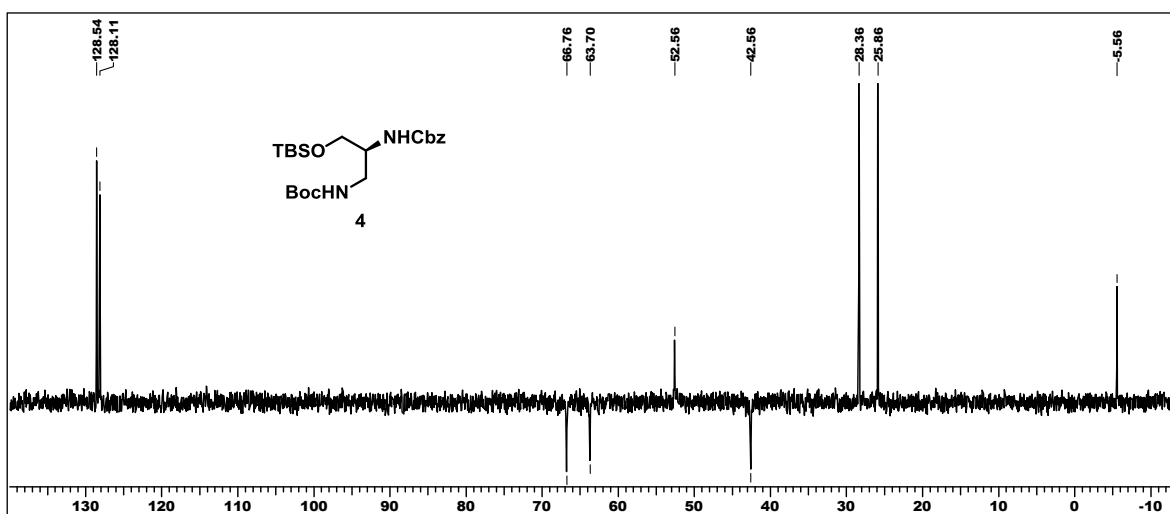
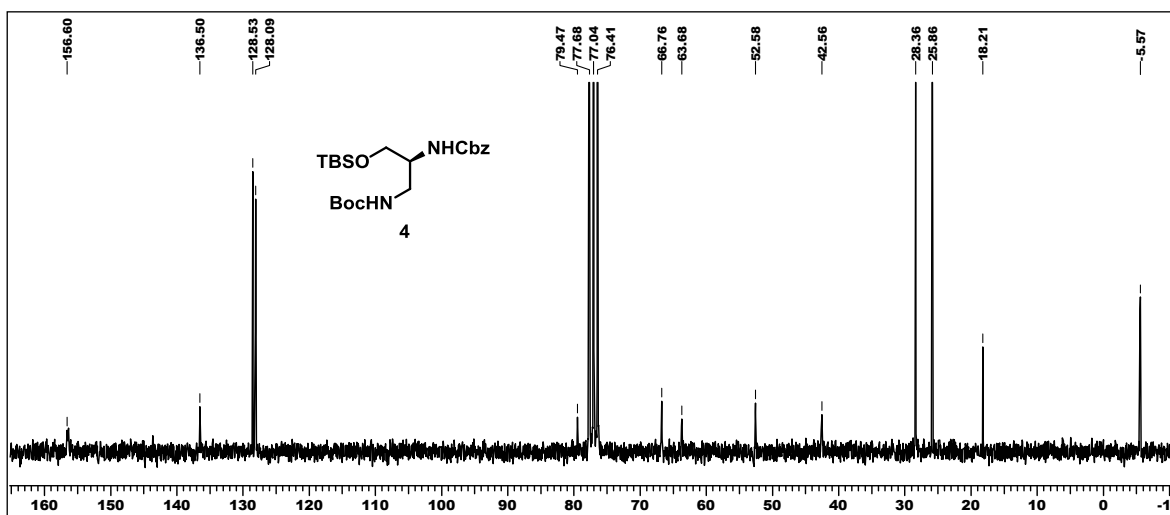




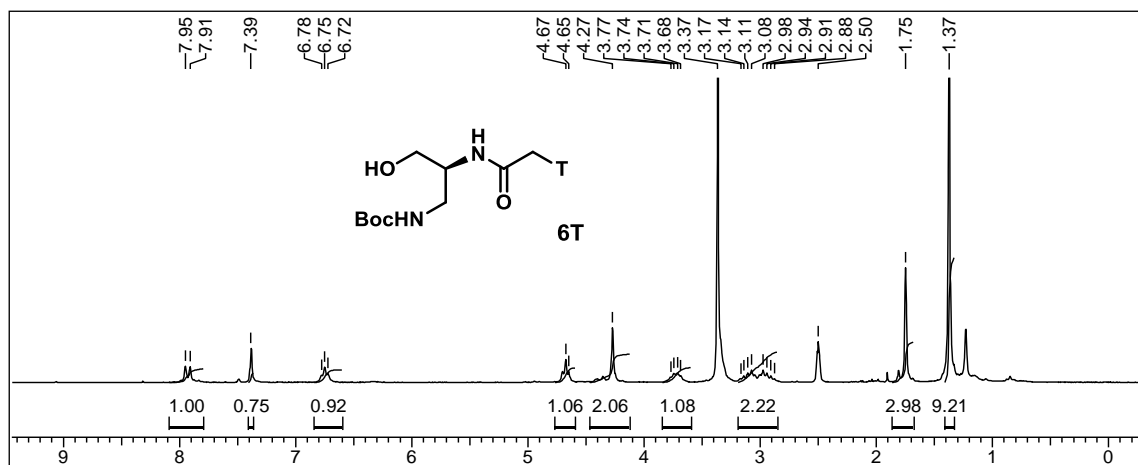
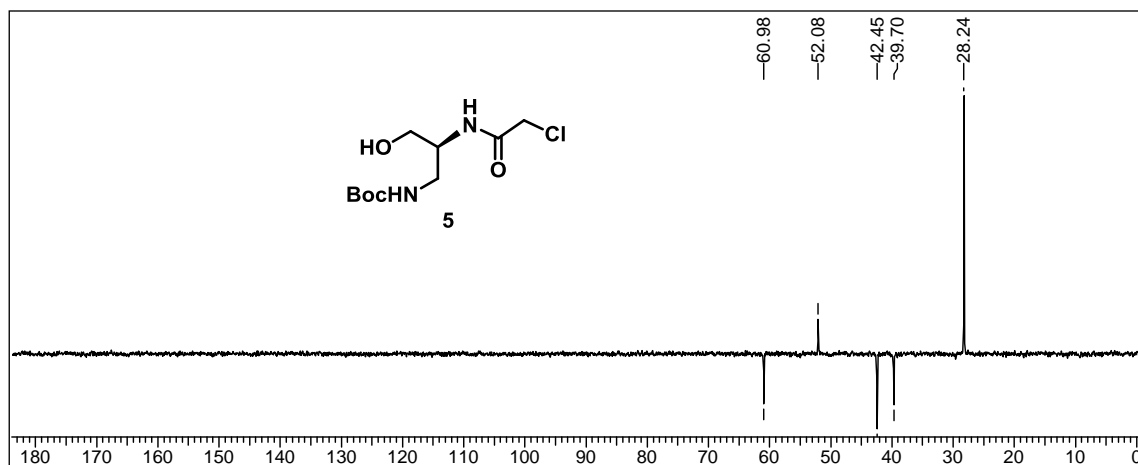
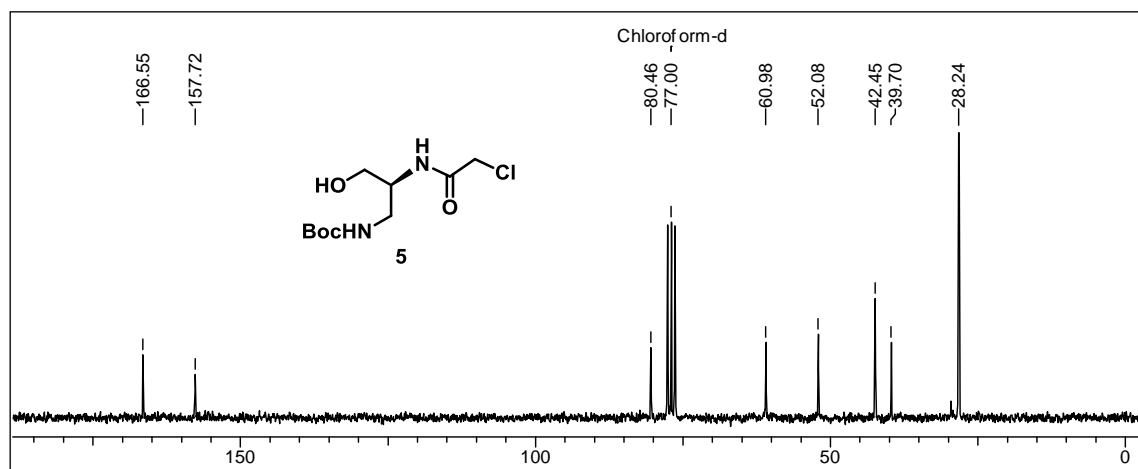




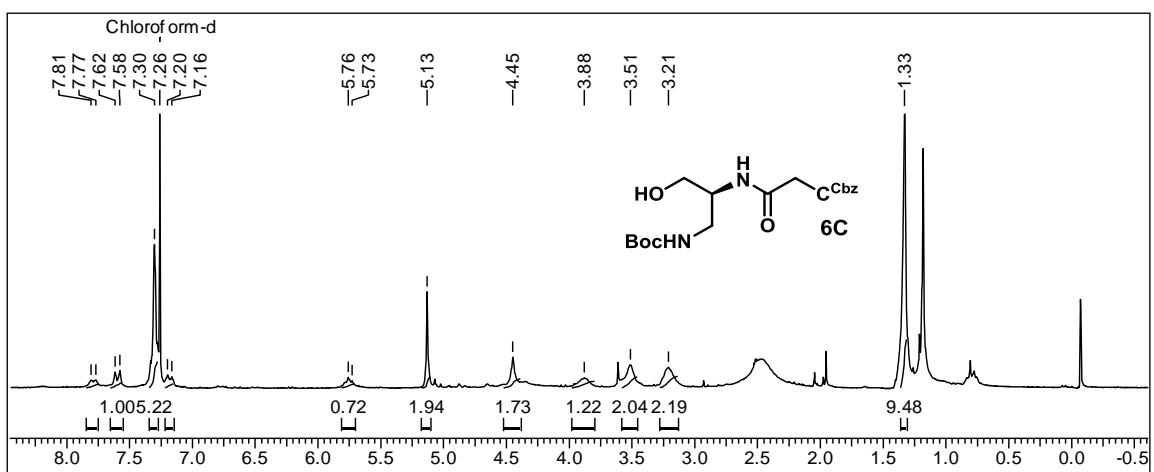
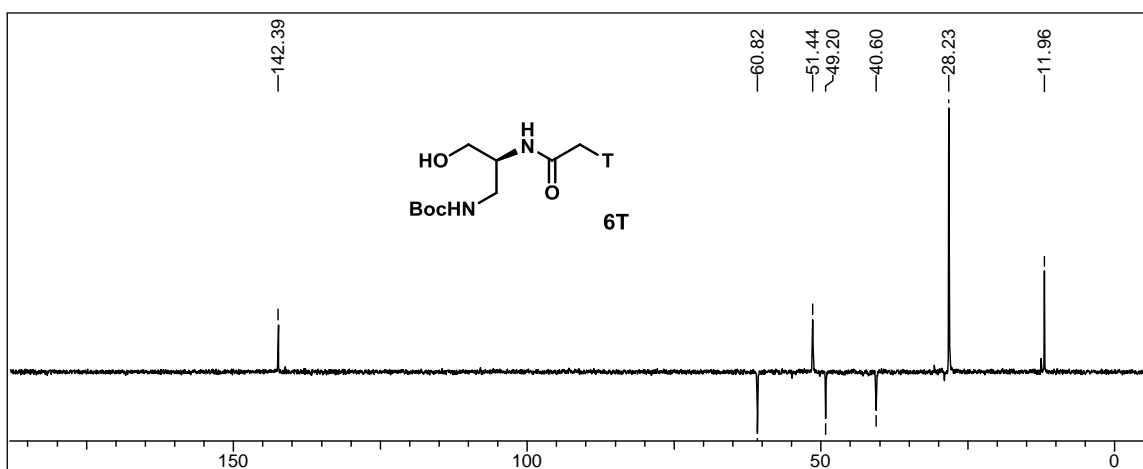
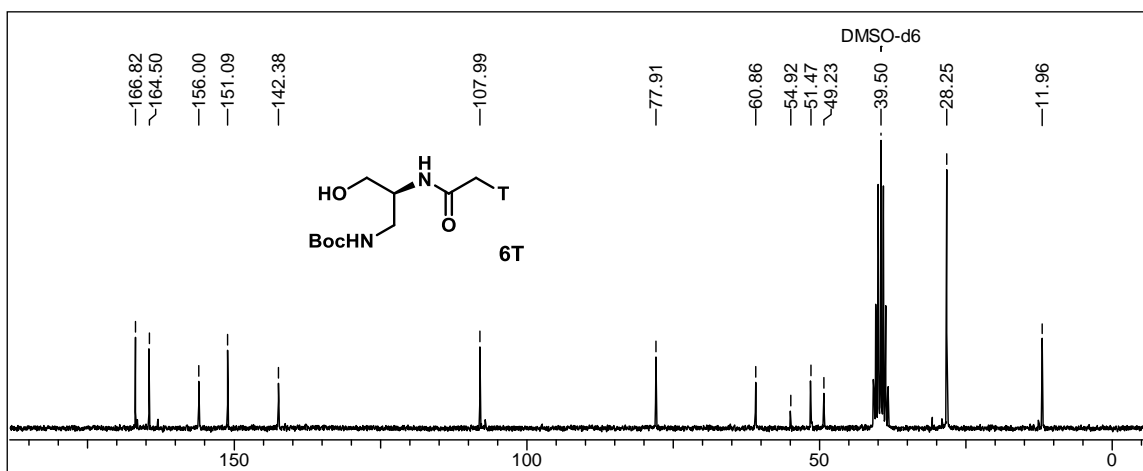
Chapter 3



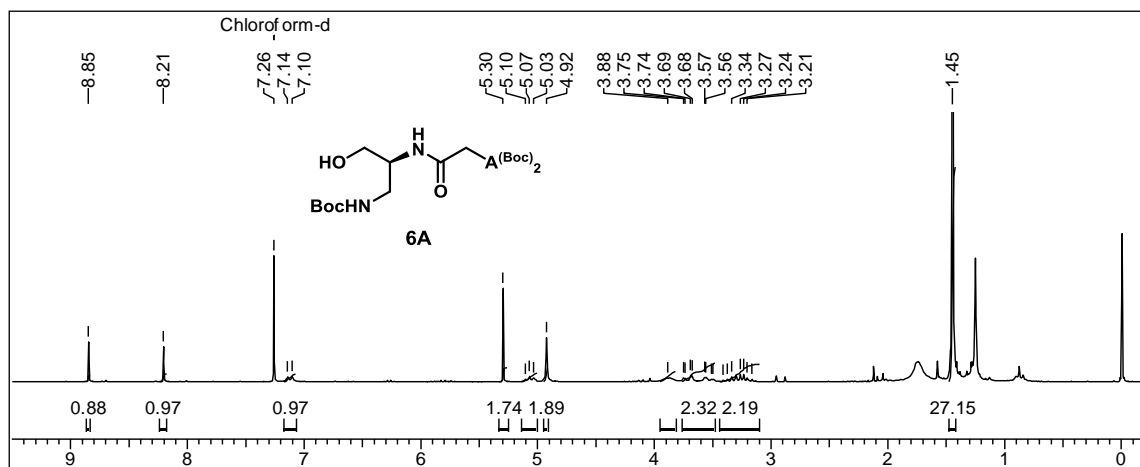
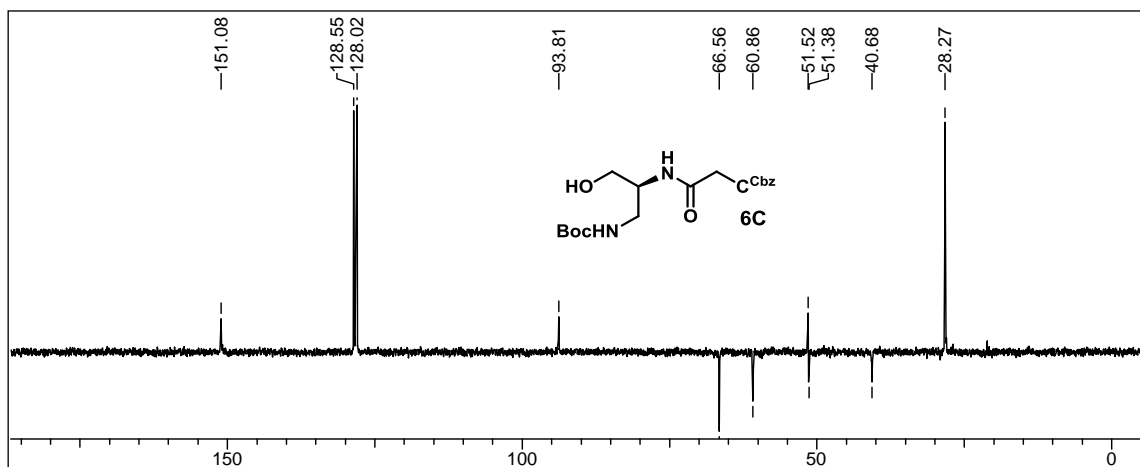
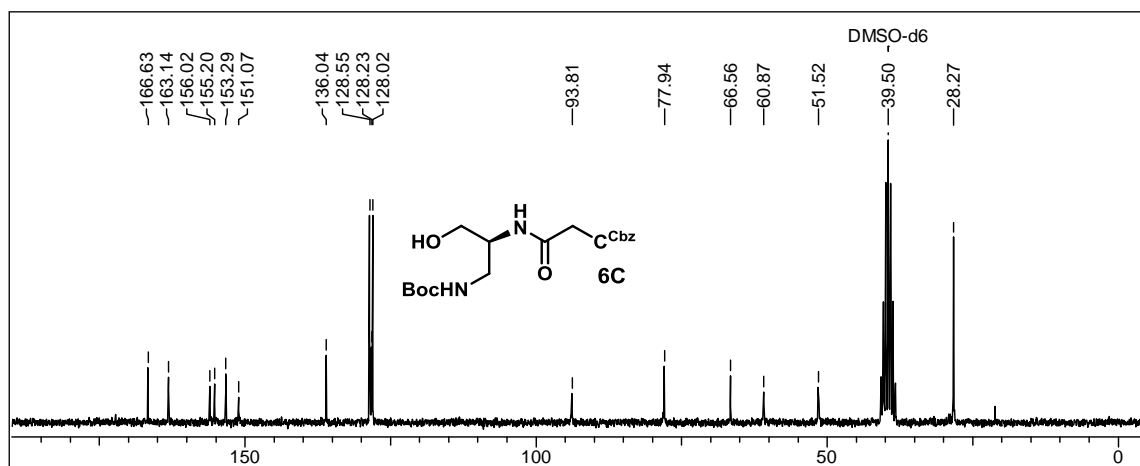
Chapter 3

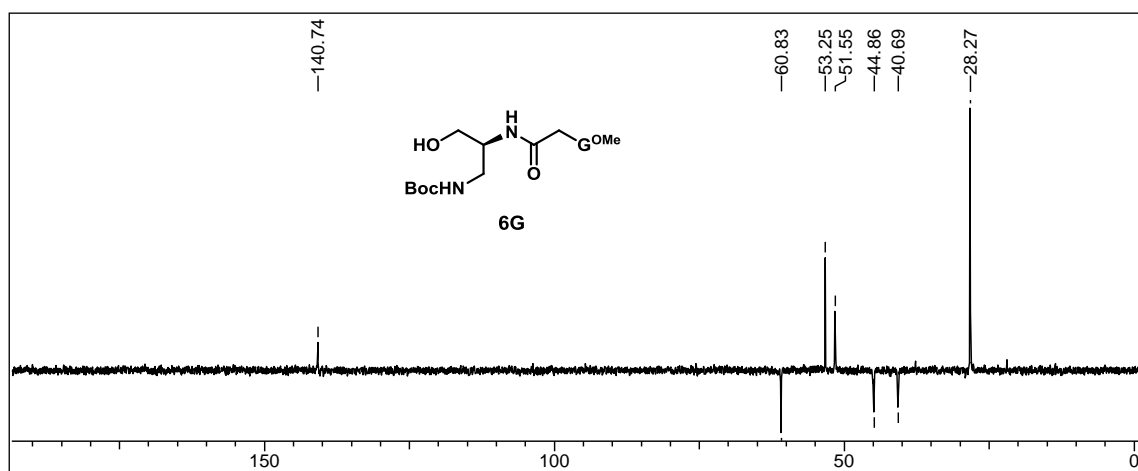
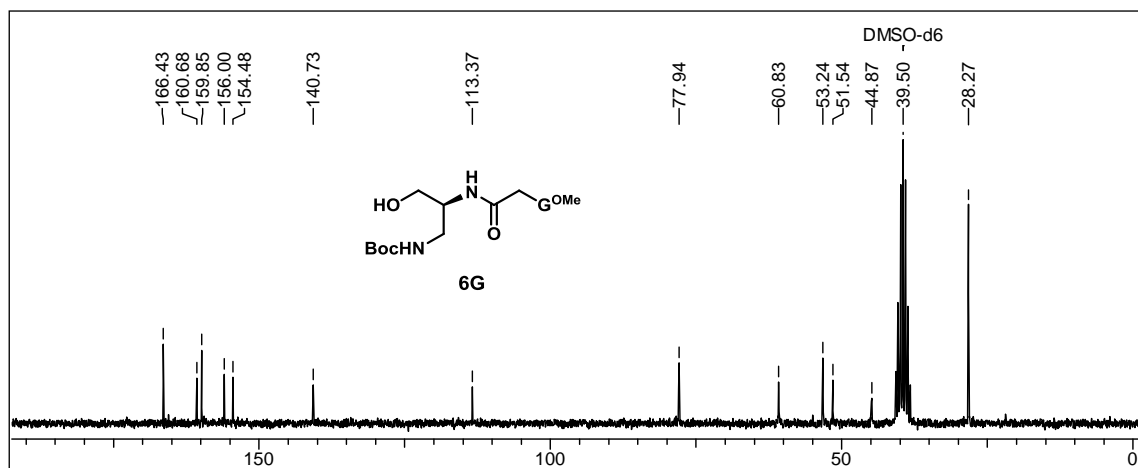
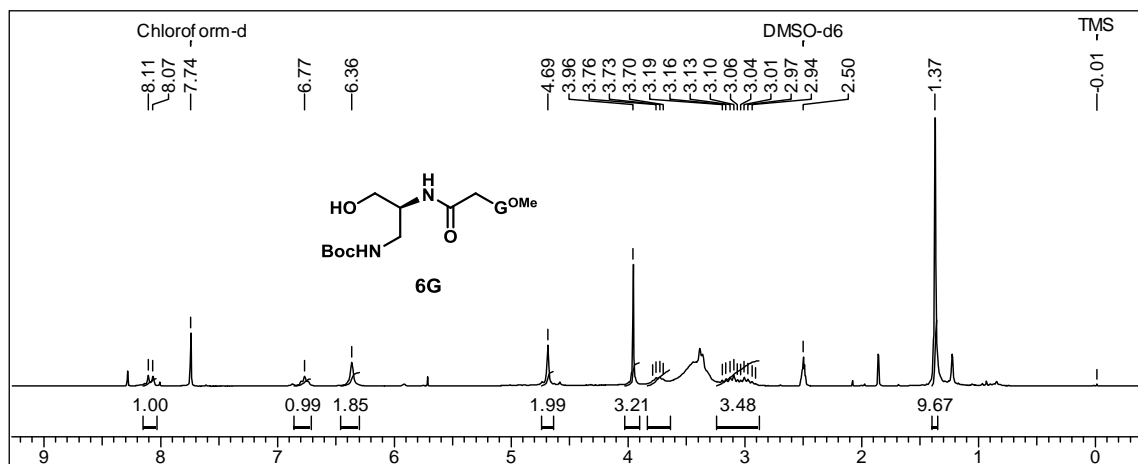


Chapter 3

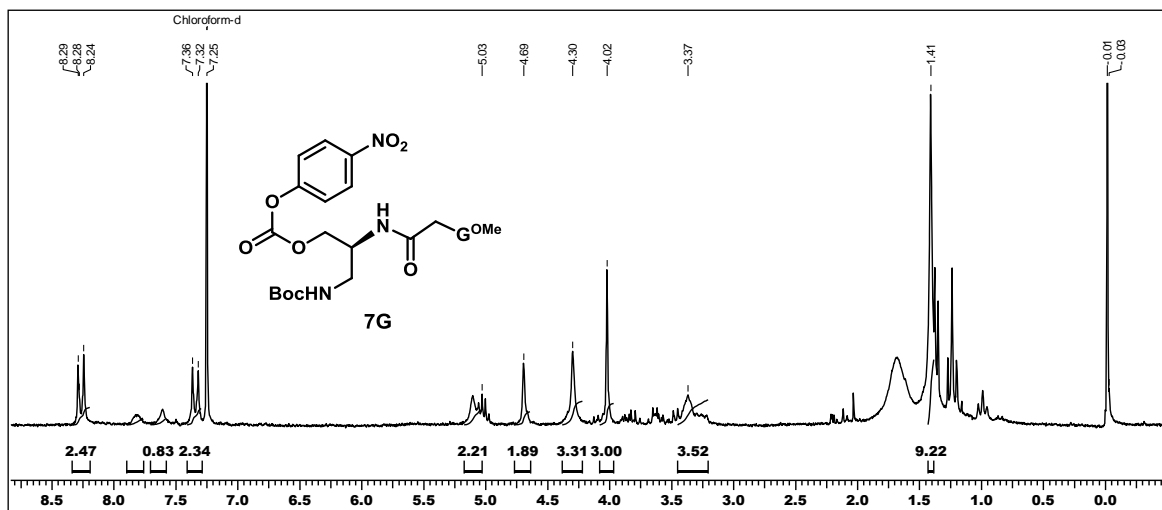
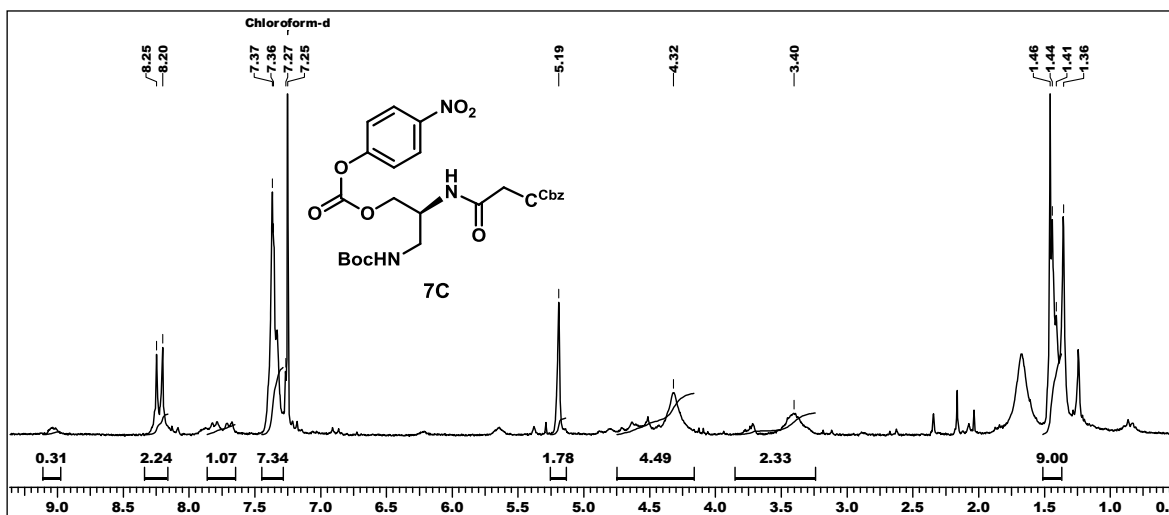
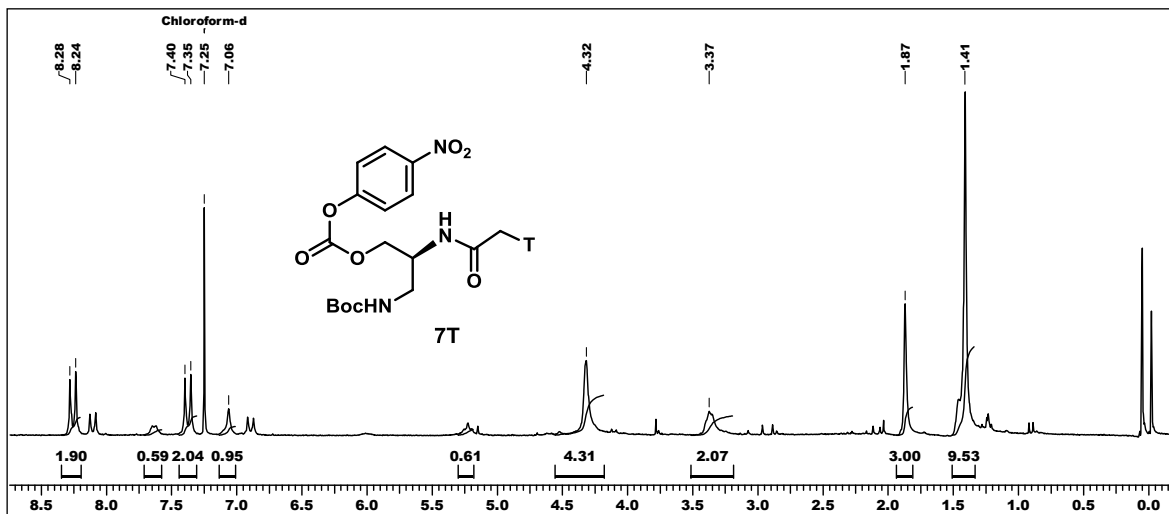


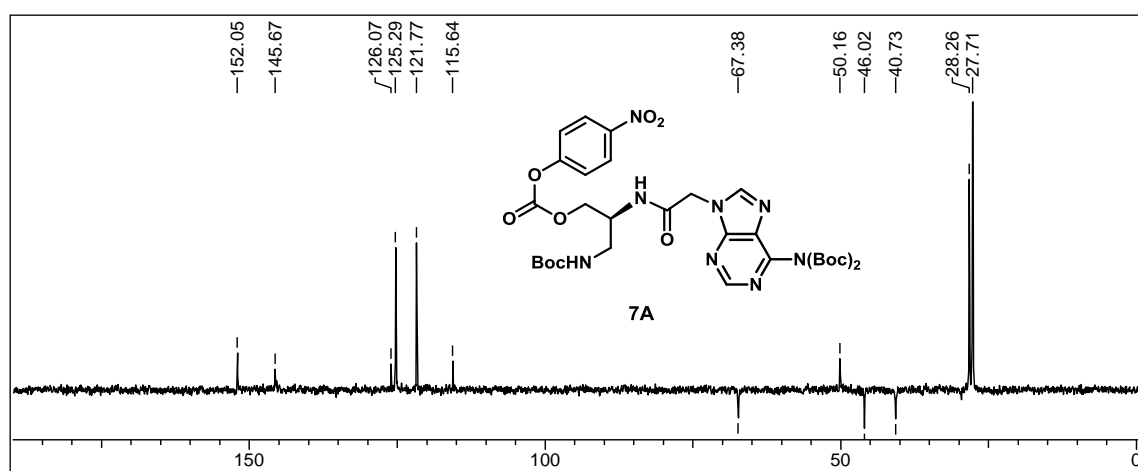
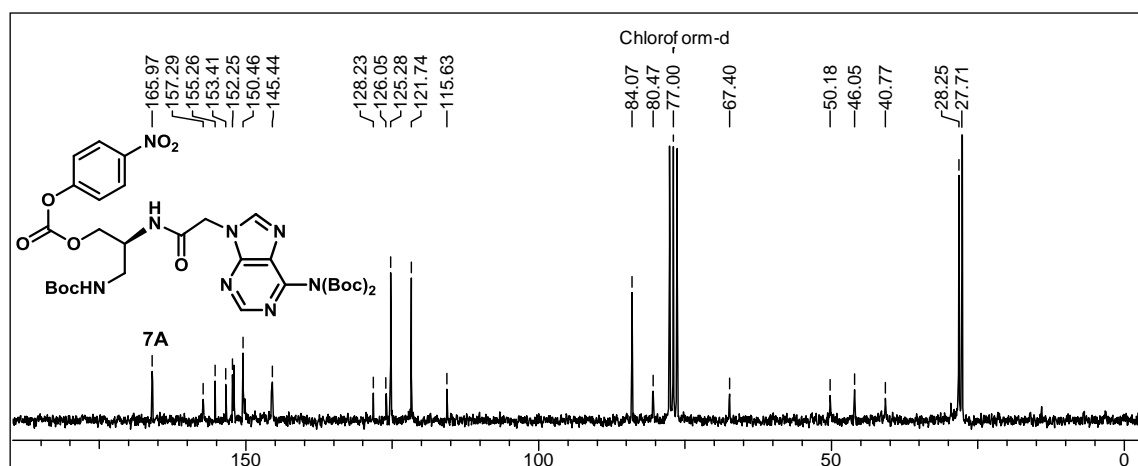
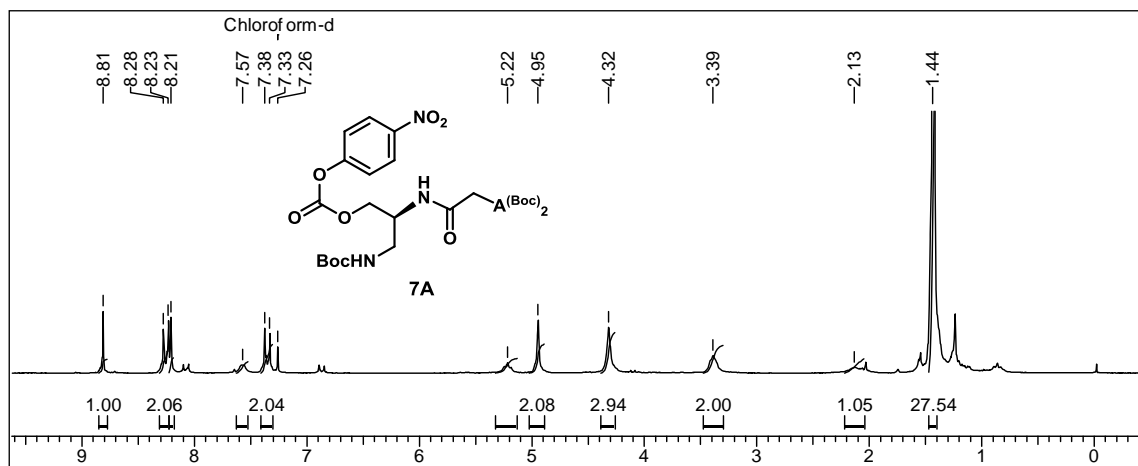
Chapter 3



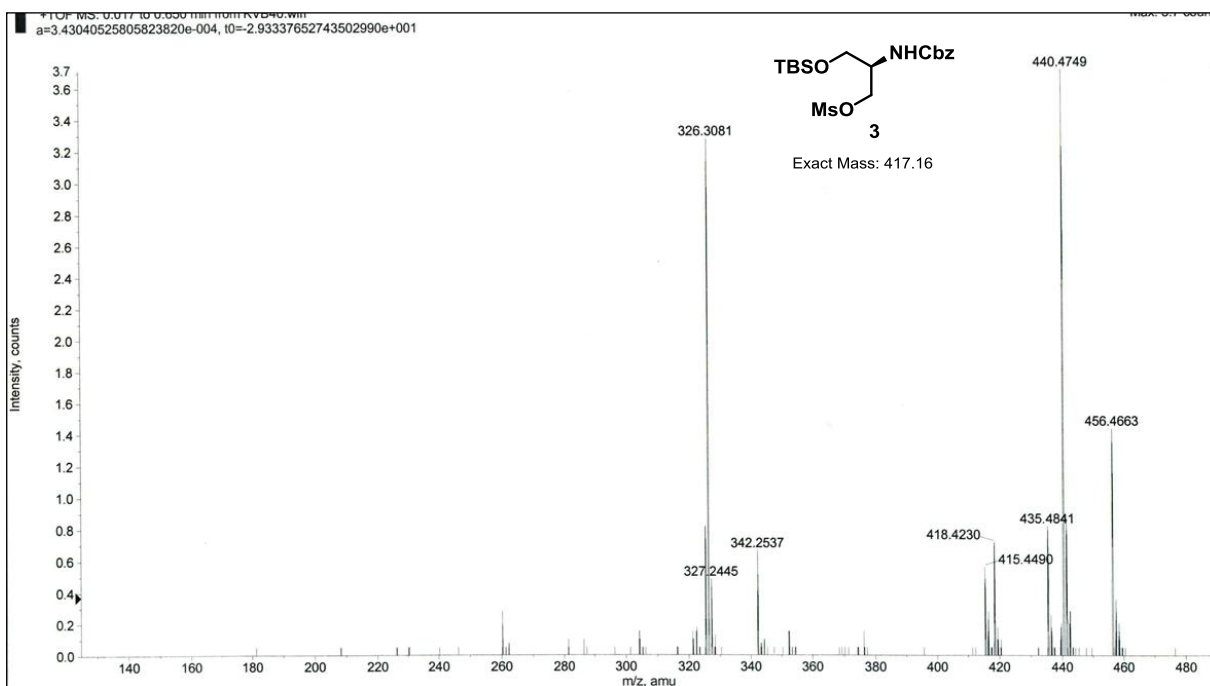
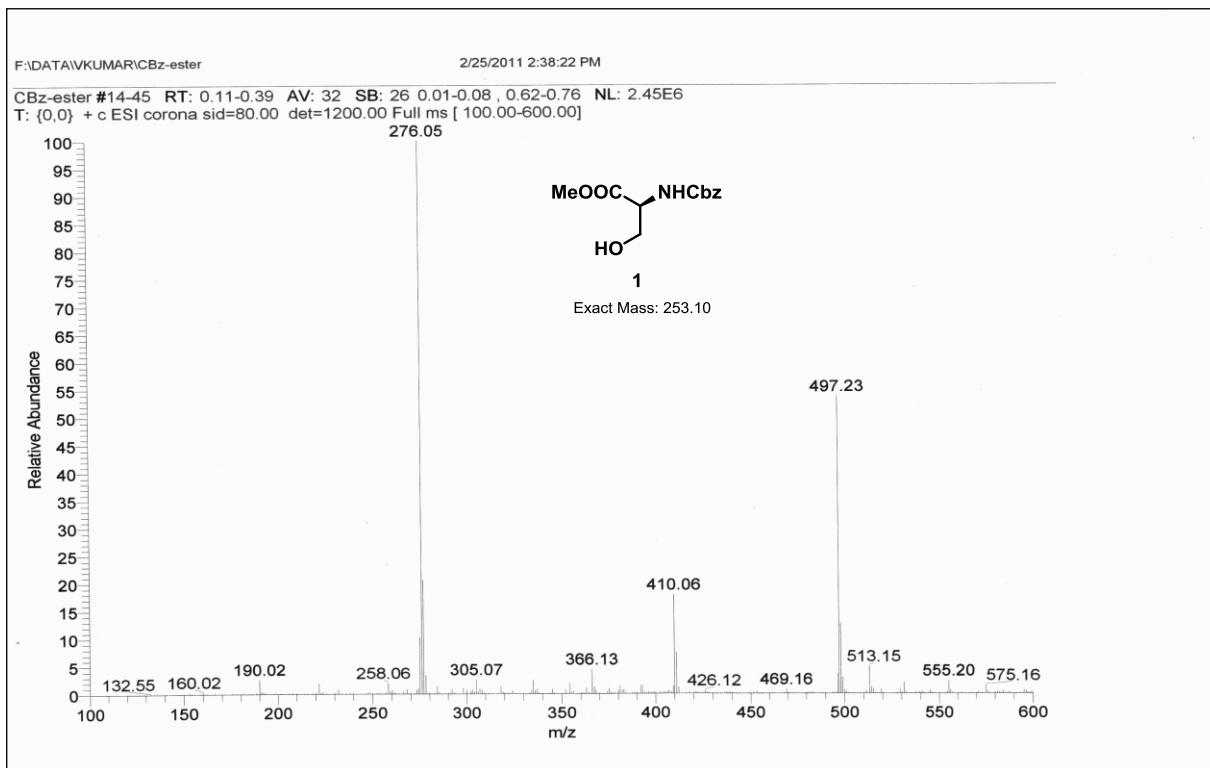


Chapter 3

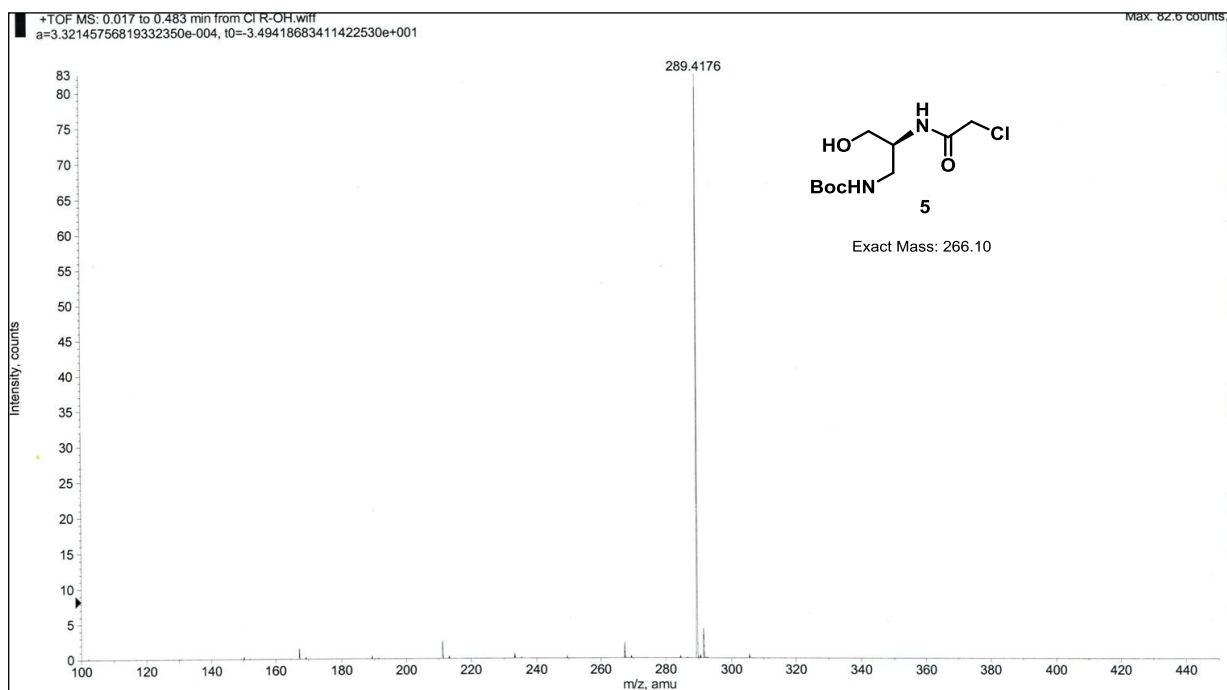
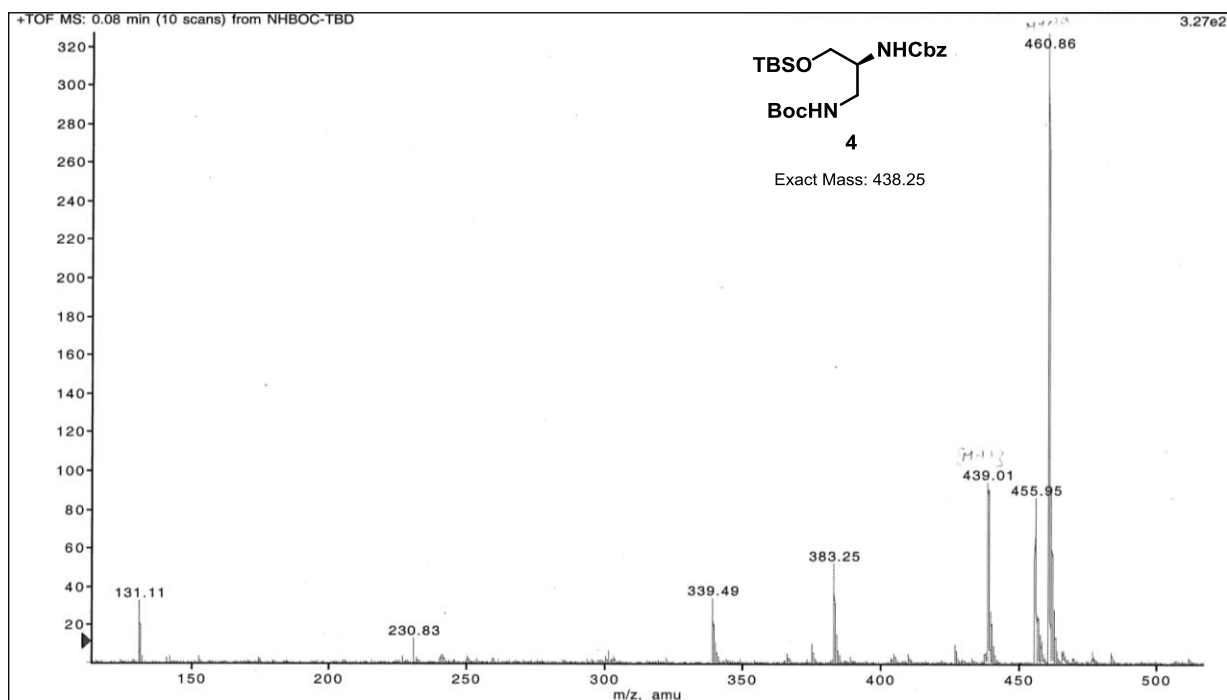




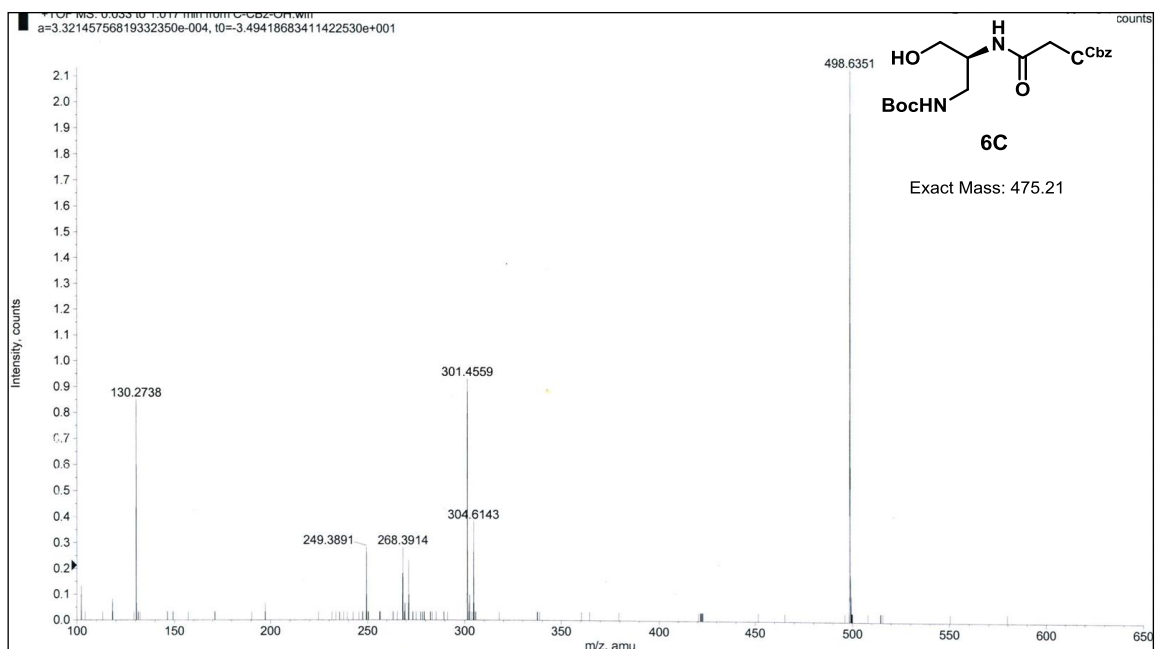
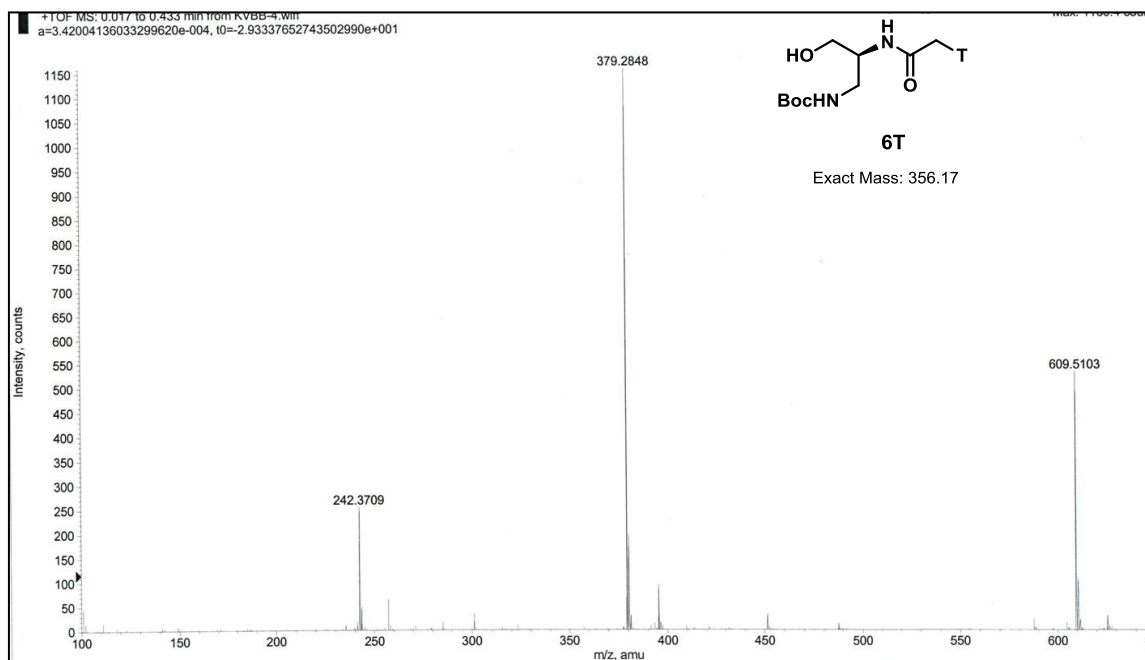
Chapter 3



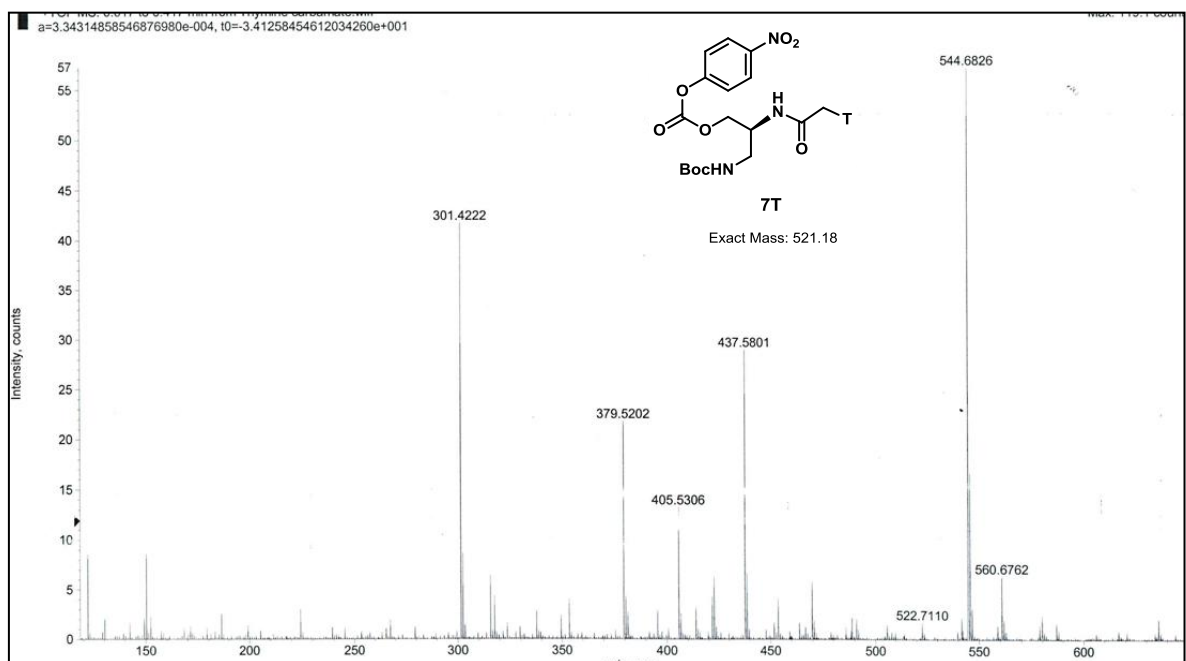
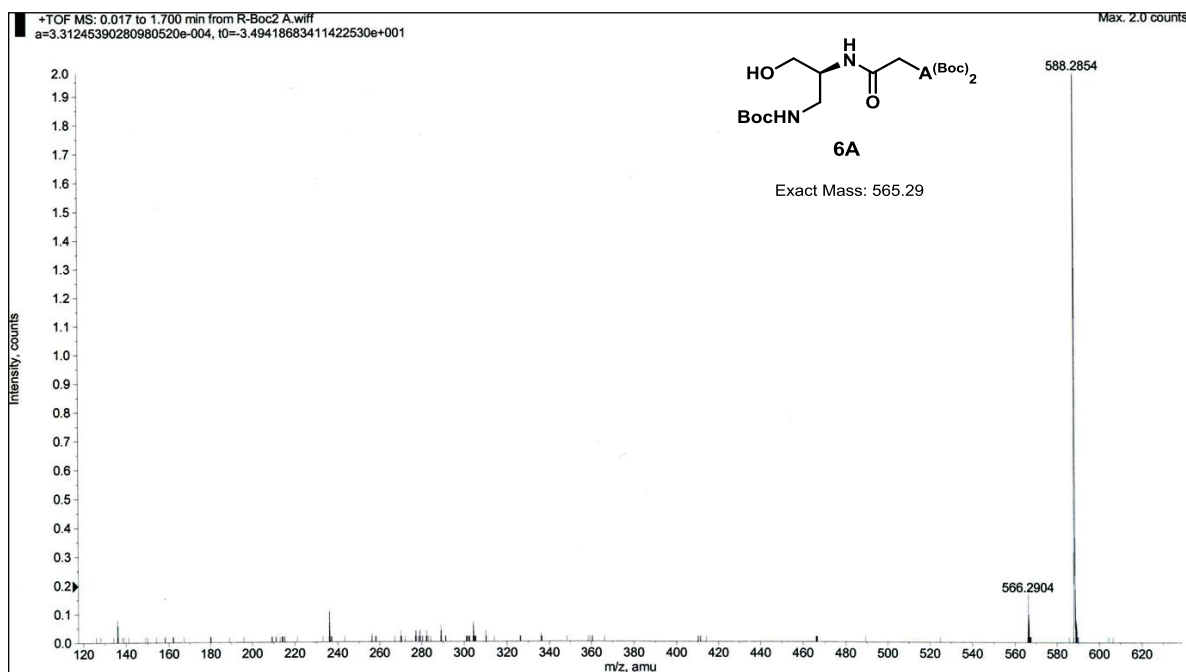
Chapter 3



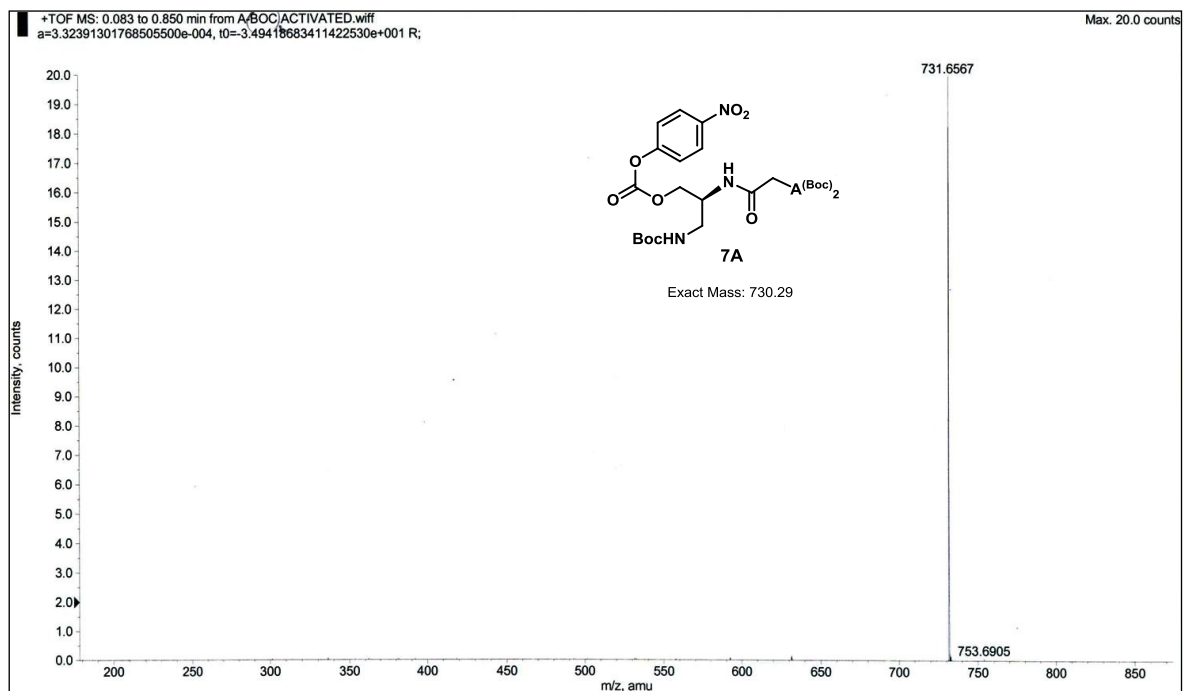
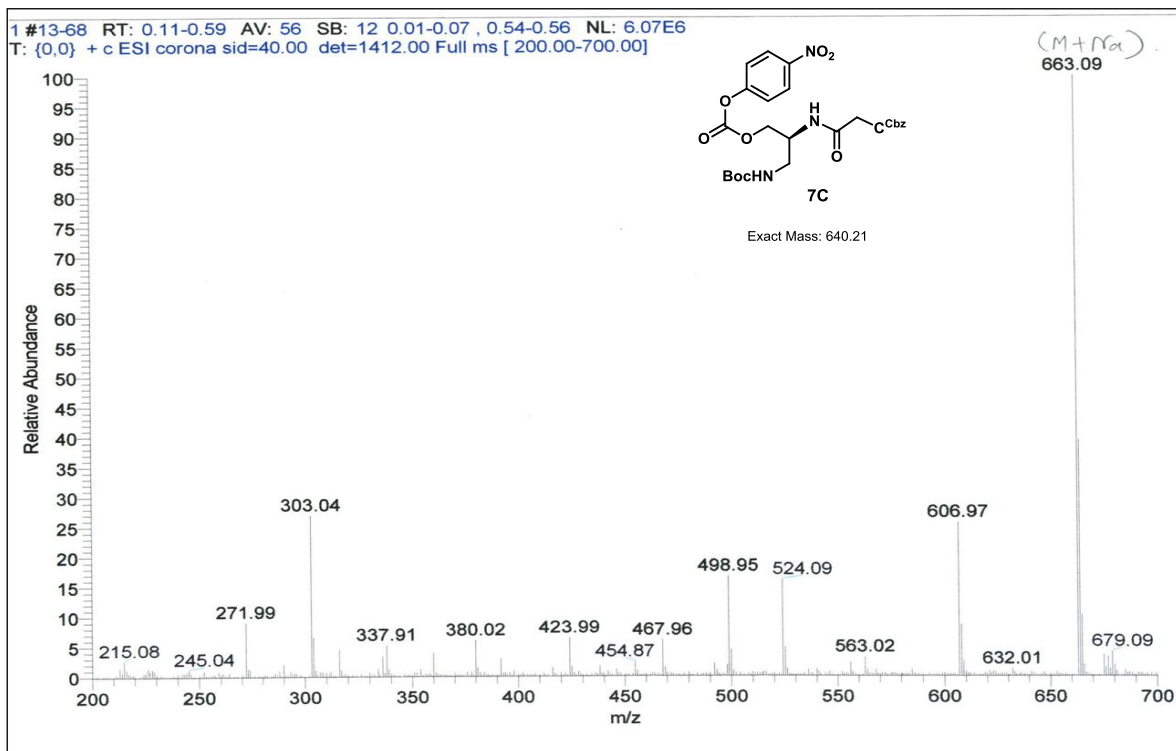
Chapter 3



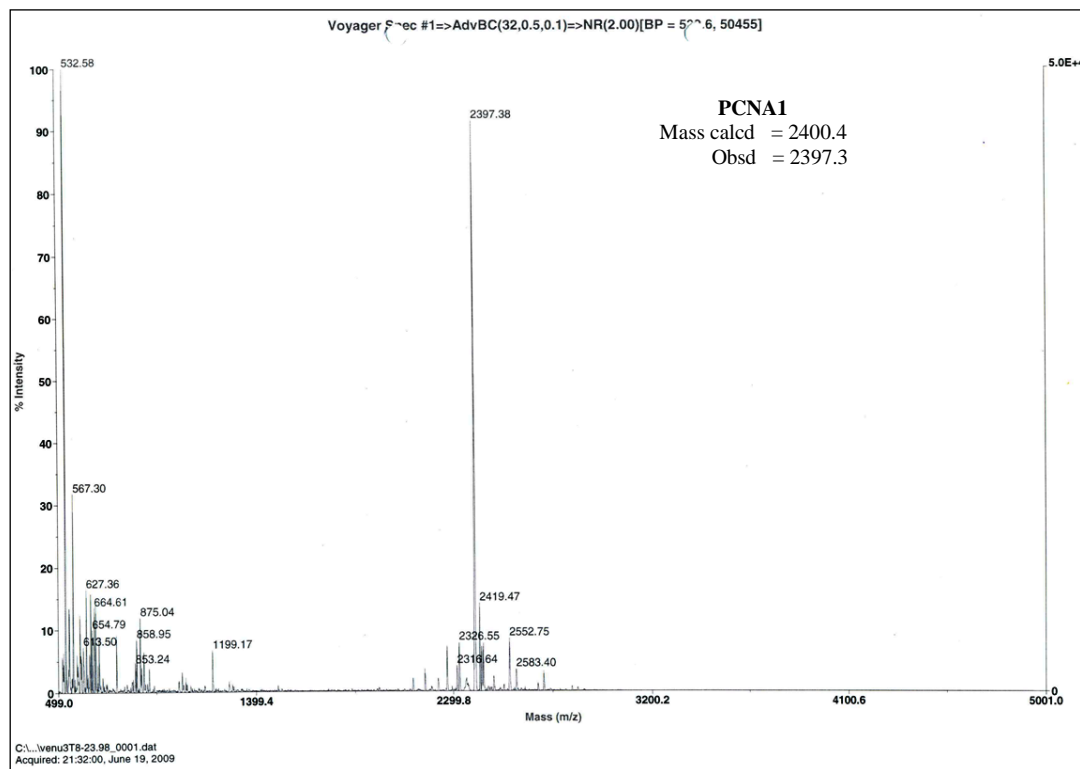
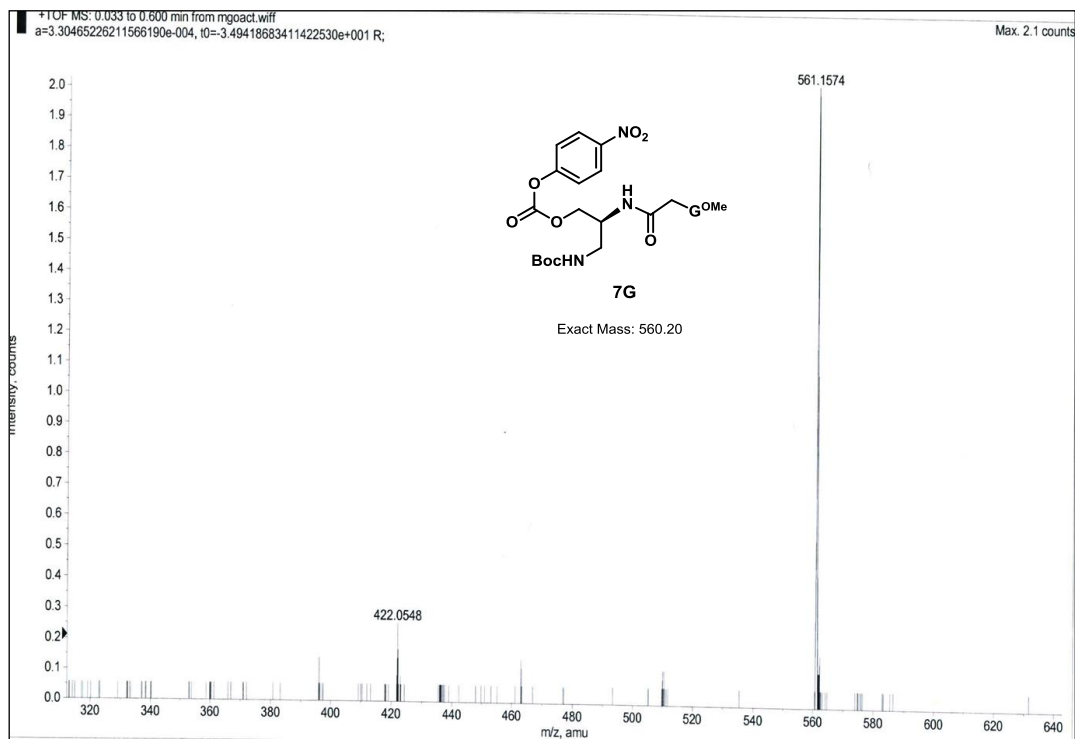
Chapter 3



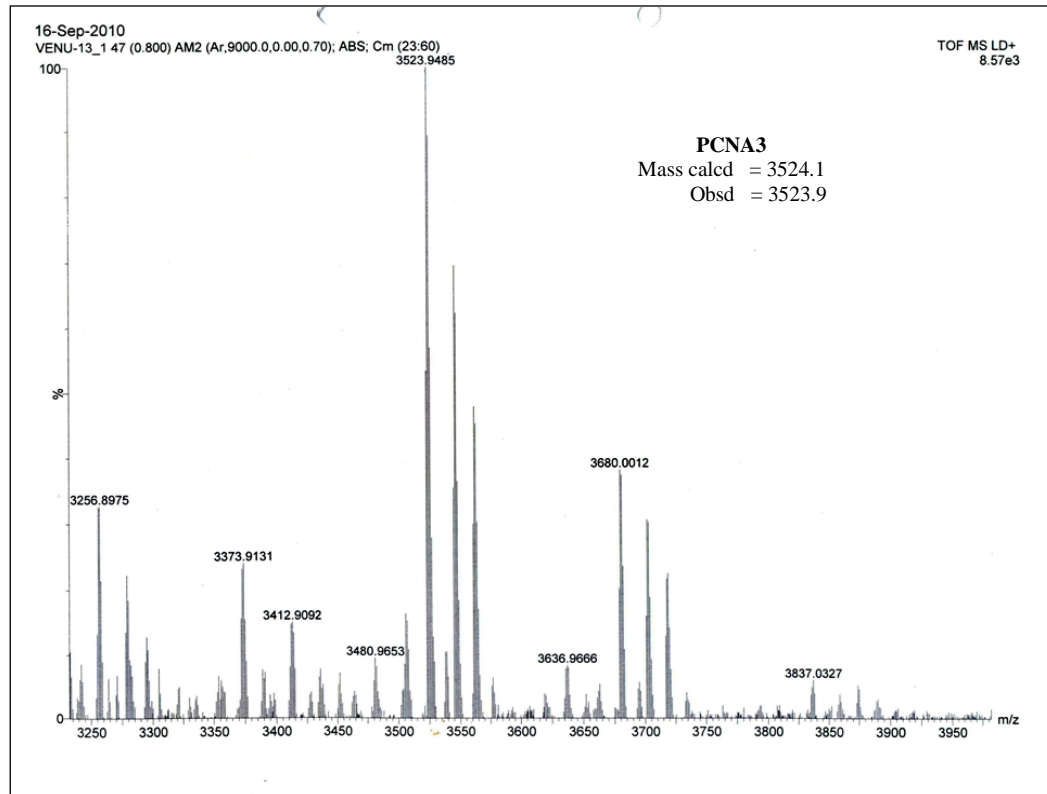
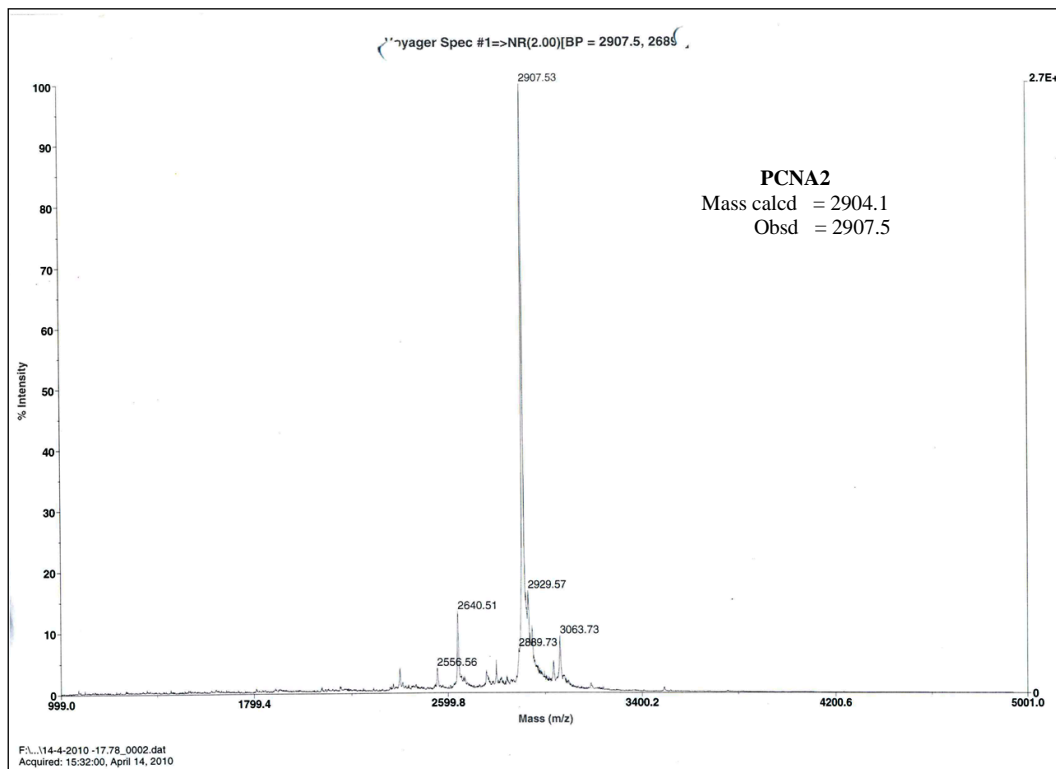
Chapter 3



Chapter 3



Chapter 3



Section 3B: Evaluation of different backboned- and 2'-sugar-modified nucleic acids for DNA diagnostics in conjunction with GO

3B.1 Introduction

Successes in genome sequencing²³ have tremendously impacted diagnostics in recent times.²⁴ Another most important discovery has been that of graphene²⁵, a novel single-atom-thick and two-dimensional carbon material with extraordinary electronic, mechanical and thermal properties and the dispersible functionalized graphene i.e. graphene oxide (GO)²⁶ which can be easily synthesized by chemists in the laboratory with a high degree of reproducibility.²⁷ Incidentally, amalgamation of these two discoveries has also started attracting attention in the last couple of years and some techniques are being developed that would greatly simplify the analysis of biological samples.²⁸

Sensitive, selective, rapid, and cost-effective analysis of biomolecules is important in clinical diagnostics. Carbon nanostructures, such as carbon nanotubes,²⁹ carbon nanodots³⁰ and carbon nanofibers³¹ have been used for this purpose. Tan and co-workers demonstrated the homogeneous detection of biomolecules using carbon nanotubes and single-stranded DNA (ss-DNA) assembly.³²

Recent research activities are directed towards graphene/ graphene oxide (GO) as a new platform for the sensitive and selective detection of biomolecules. In 2008, the work from Mohanty *et al.* on the fabrication and functioning of a chemically modified graphene-based label-free DNA sensor and bacterial DNA/protein chemical transistor has opened the new countenance of the graphene/graphene oxide and nucleic acids interactions.³³ Based on the reports of strong noncovalent binding of GO with nucleobases and aromatic compounds,³⁴ Lu and co-workers anticipated that GO could bind dye-labeled ssDNA and quench the fluorescence of the dye (Figure 13).³⁵ Also, in the presence of a target sequence, because of specific interactions between the nucleobases of the target and the labelled DNA in addition to the π - π stacking between the nucleobase pairs, the interactions between the dye-labeled DNA and GO get disturbed, releasing the dye-labeled DNA from the GO, resulting in restoration of dye fluorescence (Figure 13). They successfully demonstrated

that GO, like carbon nanotubes, can be used as a platform for fast, sensitive, and selective detection of biomolecules.³⁵

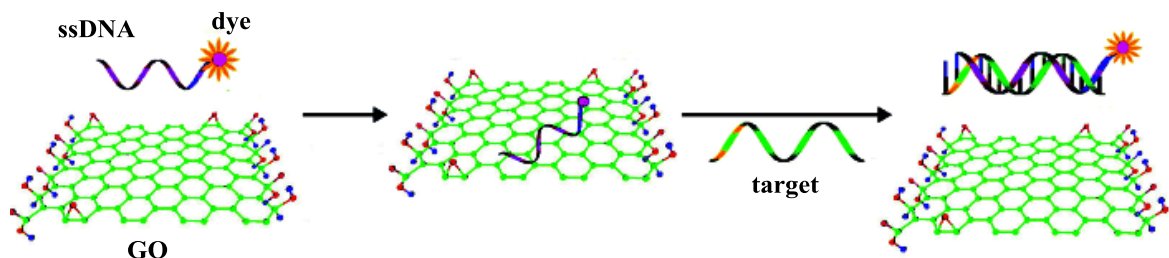


Figure 13 Schematic representation of the target-induced fluorescence change of the ssDNA-FAM-GO complex (Figure adopted from ref. ³⁵)

He and co-workers reported the graphene oxide (GO)-based multicolor fluorescent DNA nanoprobe that allows rapid, sensitive, and selective detection of DNA targets in homogeneous solution by exploiting interactions between GO and DNA molecule.³⁶ Wu and co-workers reported several factors like salt, pH, solvent and temperature that affect the adsorption and desorption of 12-, 18-, 24-, and 36-mer ss-DNA from graphene oxide surface.³⁷ They have demonstrated that the shorter DNAs are adsorbed more rapidly and bind more tightly to the surface of graphene. The adsorption is favored by a lower pH and a higher ionic strength. The presence of organic solvents such as ethanol can either increase or decrease adsorption depending on the ionic strength of the solution. They have also shown that the adsorbed DNA can also be exchanged in a non-specific way by the free DNA in solution.³⁷

Wu *et al.* reported a fluorescent biosensor for sequence-specific recognition of ds-DNA with a detection limit of 14.3 nM, developed based upon the DNA hybridization between dye-labeled ss-DNA and ds-DNA.^{28c} Guo *et al.* demonstrated that hemin-graphene hybrid nanosheets (H-GNs) retained the intrinsic property of peroxidase-like activity of hemin and also their ability to differentiate ss- and ds-DNA. They developed a label-free colorimetric detection system for single-nucleotide polymorphisms (SNPs) in disease-associated DNA.³⁸ As most of the reported fluorescent sensors relied on DNA probes adsorbed by GO, which might suffer from nonspecific probe displacement and false

positive signal, Haung and Liu reported a molecular beacon (MB) comprised of a DNA-fluorophore and covalently connected to GO-quencher by a DNA hairpin, with a detection limit of 150 pM.³⁹ The biomedical applications evaluated based on the graphene/GO interactions with the nucleic acids include apoptosis diagnosis,⁴⁰ a facile RNA extraction method,⁴¹ the quantitative detection of target miRNA expression levels in living cells,⁴² the selective and in situ detection of multiple nucleotides (like ATP, GTP and other biomolecules) in living cells,⁴³ separation of short ss- and ds-DNAs⁴⁴ and finally highly sensitive & specific sensing platform for the picomolar detection of DNA.⁴⁵

Apart from the graphene based nucleic acid detection probes, there were several detection probes that are derived from the nucleic acids itself in combination with the enzymes or other materials. The nucleic acid chemists working towards therapeutics have developed numerous chemically modified nucleic acids which exhibit strong and sequence specific hybridization properties with the complementary sequence.⁴⁶ Out of the large pool of synthetic nucleic acids that are developed for antisense technologies, very few have been evaluating for nucleic acid-diagnostic applications. Because of their hallmark ability of strong duplex formation, peptide nucleic acids (PNAs)^{1a} and locked nucleic acids (LNAs) are being explored for diagnostic applications, the former being dominated as they are highly resistant to degradation by nucleases and proteases than the later.

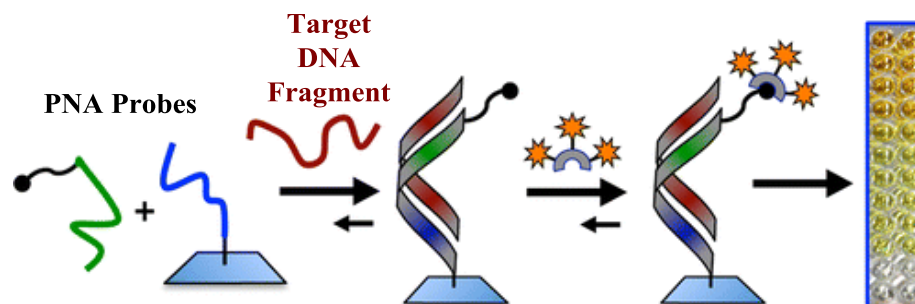


Figure 14 Schematic of qualitative and quantitative detection of anthrax DNA by using PNA-derivative derived probe (adopted from ref⁴⁷).

Recently, Micklitsch *et al.* developed chemically modified PNA-based, enzyme-linked reporter assay and demonstrated excellent detection levels of anthrax DNA qualitatively and quantitatively (Figure 14).⁴⁷ There are several other detection assays

which were developed based on the unique properties of PNAs. On the other hand, locked nucleic acid-based probes also gain interest in diagnostic applications. Obernosterer *et al.* developed locked nucleic acid-based probe for the *in situ* detection of miRNAs.⁴⁸ This ‘*in situ* miRNA detection’ which reported to take one week time for setting up the assay was greatly improved by the LNA-based probes using northern blotting for miRNA detection (one week vs. 4 h).⁴⁹ Time consuming and the poor sensitivity- the major disadvantages of the traditional DNA probes in northern blot technology was also addressed by the LNA-based probes.⁴⁹

3B.2 Rationale and objectives of the present work

The basic concept of all the GO-based probes take advantage of GO's preferential interaction with ss-DNA over ds-DNA. The exposed nucleobases in ss-DNA are adsorbed strongly with GO surface in contrast to the nucleobases in ds-DNA. The stacked pairs of nucleobases are effectively hidden in the helical structure, which prevents the direct interaction of nucleobases with GO surface. Therefore, when the fluorescently labeled ss-DNA probes bound to the GO surface hybridize with their complementary target DNA, forming a duplex, it is desorbed from the GO surface, leading to a recovery of fluorescence (Figure 15, Path A).²⁸

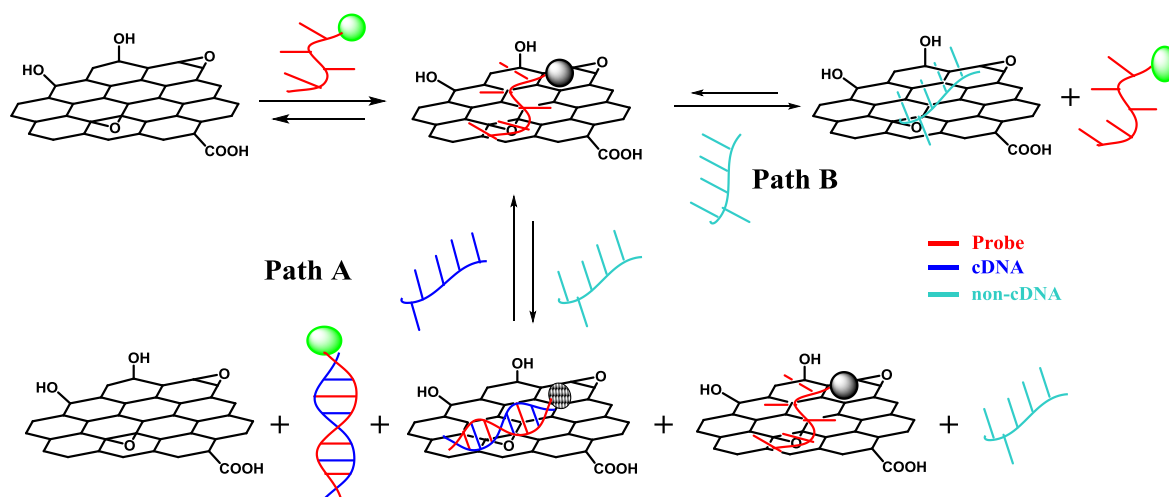


Figure 15 Schematic representation of fluorescence quenching of labeled probes with GO and the possible pathways in restoring the fluorescence by cDNA/random DNA.

The phosphodiester linkages in probe as well as target allows the reversible adsorption/desorption process and would lead to unwanted Path B shown in Figure 15. This technique of detection would be most reliable if nonspecific displacement of the probe from the GO surface by non-complementary DNA could be avoided (Figure 15, Path B). The sensitivity of the specific detection may be controlled by the balance between the interaction of GO with the probe, the strength of probe:target complex, interaction of GO with the probe:target complex, and also the interaction of target with the GO surface (Figure 15). The fluorescence intensity in duplexes also sometimes is dependent on the base sequence and in that case, the fluorescence quenching upon duplexation could be a limiting factor.⁵⁰ In addition to the π - π stacking interactions of the nucleobases with the planar GO surface, several other mechanisms have also been suggested to play a role in DNA-GO interaction, e.g. hydrogen bonding, electrostatic, van der Waals, and hydrophobic interactions.^{28a-d, 38, 51}

Considering these points, we thought that the synthetic DNA analogues and mimics with high binding affinity to DNA and with altered backbone characteristics⁴⁶ could have advantages over natural DNA for this application. Use of modified oligomers in DNA detection in conjunction with GO has not yet been reported in literature despite their strong, sequence-specific binding with natural DNA. We present our studies that evaluate the most promising known synthetic DNA analogs/mimics capable of forming strong duplexes with cDNA in comparison with R/S-PCNA of the present study as nucleic acid detection probe in conjunction with GO (Figure 15).

Besides the well established PNA- and LNA-modifications, R-PCNA¹³ earlier reported from our group and S-PCNA of the present study which exhibited even stronger binding than PNA to the cDNA and also the therapeutically important modification 2'-*O*-(*S*-2-amino-3-methoxypropyl)-substituted oligomers (described in Chapter 2) are included in this study to examine their utility in GO-mediated interactions with target DNA. Several sequences of nucleic acid analogues (Figure 16) are synthesized, characterized and evaluated in the quest for finding a better probe. The OMe modification with phosphodiester linkage are used to study the effect of 2'-alkyl substitution, and DNA/LNA mixmers are

used as they are known to form very strong duplexes with cDNA. PNA is used as achiral, uncharged backbone mimic of DNA whereas PCNA as a chiral uncharged backbone mimic of DNA.

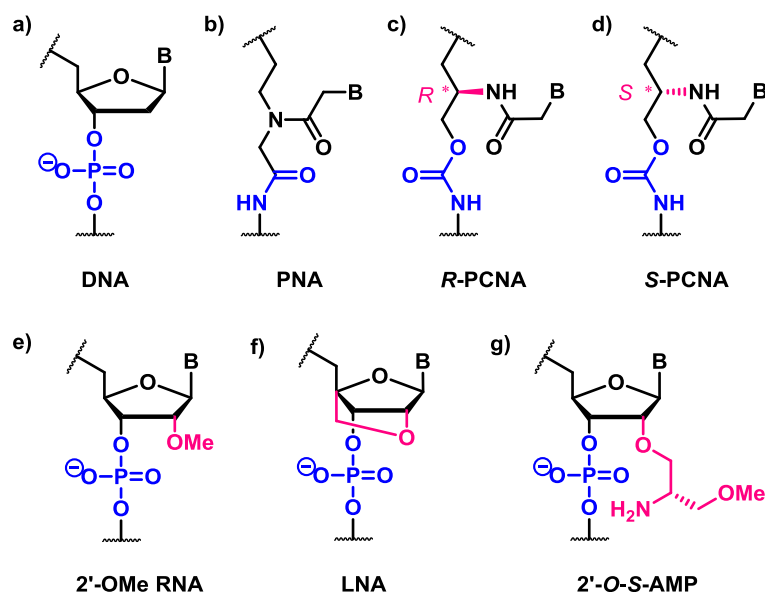


Figure 16 DNA and selected chemically modified nucleic acids that exhibit strong duplexation abilities and/or expected to have additional electrostatic interactions with the probe.

This section deals with the initial understanding of interactions of different backbone octathymine probes [namely phosphate- (DNA), polyamide- (PNA) and polycarbamate-(PCNA), Figure 16a-16d] with GO using fluorescence quenching and CD studies and also the fluorescence restoration with cDNA. The similar studies were extended to mixed pyrimidine/purine sequences of 10-mer and 12-mer length. Change of pH for select sequences and the thermal desorption studies of the probes from GO surface are also performed to correlate the results to the cDNA-induced desorption. Further we have studied the effect of sequence length on the desorption of modified DNA duplexes (Figure 16e-16g).

3B.3 Synthesis of the materials used

3B.3.1 Synthesis of graphene oxide (GO)

Graphene oxide (GO) was prepared from graphite powder according to modified Hummers' methods. Graphite powder was treated with strong oxidising media of H_2SO_4 and KMnO_4 at 98°C for 15 min and followed by the addition of 30% H_2O_2 . The yellow suspension was separated by centrifugation and washed repeatedly with dil. HCl. The final product was dried over phosphorus pentoxide *in vacuo* for two days.

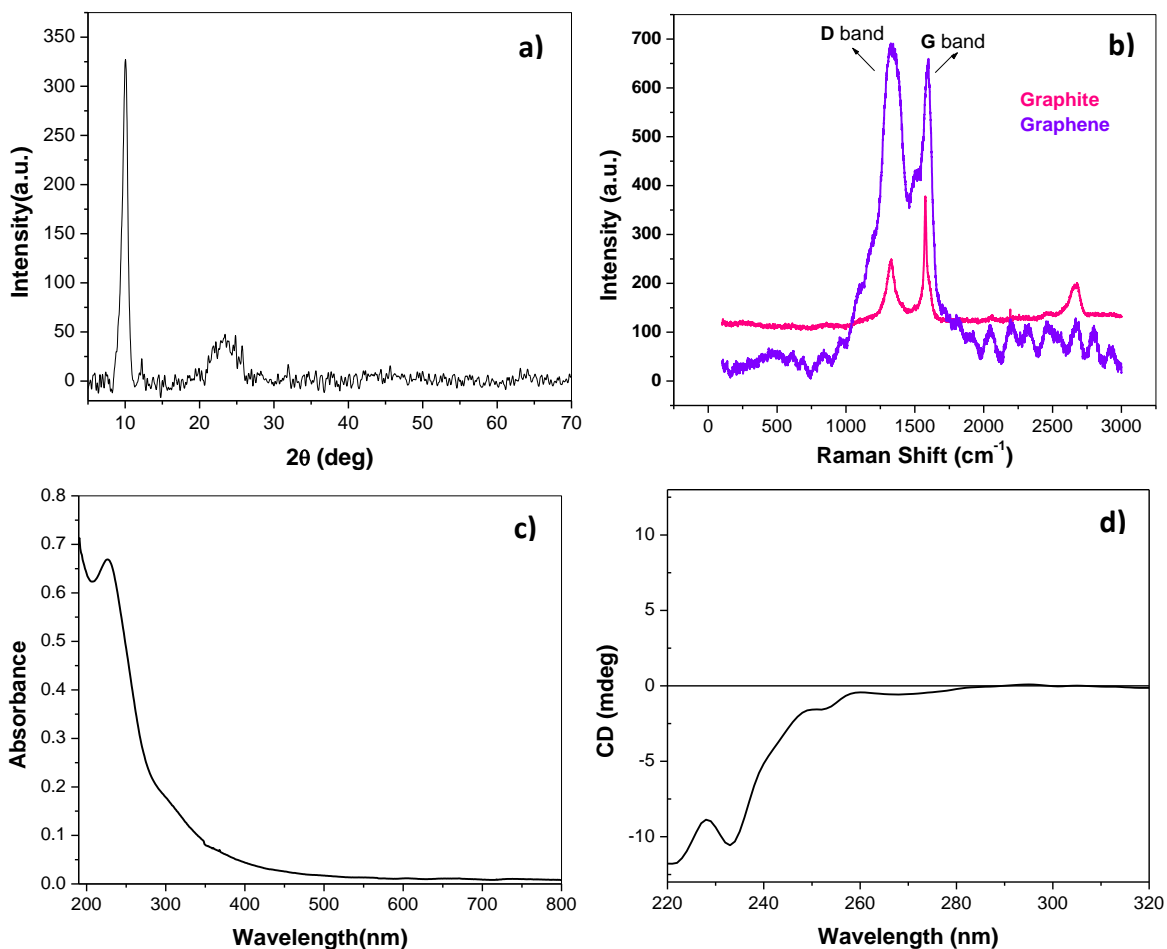


Figure 17 Characterization of GO: **a)** XRD spectrum of graphene oxide and **b)** Superimposed Raman spectra of graphite (pink) and graphene oxide (violet) **c)** UV absorbance scan of GO and **d)** CD spectrum of GO.

The prepared GO was characterized by XRD,⁵² Raman⁵³ and UV spectroscopy, (Figure 17). The XRD shows a sharp peak (2θ) at 11.3° , a characteristic peak of GO and a

minor peak at 24° indicating the partially oxidized/unoxidized graphite (Figure 17a). Raman studies of the material (Figure 17b) showed a significant drop in the G-band intensity of oxidized material, ensuing from the breaking of the graphitic structure during oxidation. A good increase in the *ID* to *IG* ratio from 0.67 to 1.04 was observed. Also, the UV absorbance scan of the material shows a maxima at 225nm, which is a characteristic of GO (Figure 17c). In addition, the CD of the prepared GO was also recorded and it shows the characteristic negative CD signature at 232 nm (Figure 17d).

The synthesized GO used in these experiments is highly functionalized having oxidized functionalities (such as hydroxyl, carboxylates or epoxides) and aromatic/hydrophobic surfaces. Thus, there are possibilities that the binding of different backbone oligomers to the GO surface would involve different types of interactions such as electrostatic and hydrogen-bonding with different backbones (phosphate, amide, carbamate), in addition to their hydrophobic π - π base-stacking interactions with GO, common to all the sequences.

3B.3.2 Synthesis of labelled DNA mimic/analogue probe and MALDI-TOF characterization

The synthesis of enantiomeric *R*-PCNA¹³ and *S*-PCNA monomers were achieved from the same chiral source L-serine (Figure 18 as discussed in Chapter 3, Section A), and the oligocarbamates were synthesized manually on solid support using ‘Boc chemistry’. The polyamide backbone PNA probes were also synthesized manually on solid support using -Boc chemistry by following earlier published protocols.⁵⁴ The synthesis of LNA-thymidine amidite from glucose diacetonide (Scheme 1)⁵⁵ was achieved by following the reported protocols and 2'-*O*-(*S*-2-amino-3-methoxypropyl)-substituted uridine amidite was synthesized as shown in Scheme 2 and 3, Chapter 2. The natural and modified phosphate backbone oligonucleotide probes were synthesized on a Bioautomation MM-4 DNA synthesizer using standard conditions. All the oligomeric sequences were labelled on solid support using carboxyfluorescein labelling at 5'-/N-terminus of the oligomers, purified by RP-HPLC and characterized by MALDI-TOF (Table 3). The UV- T_m studies were

performed to estimate the strength of the duplex with the complementary DNA and the data is tabulated at the respective discussion (Table 4, 5 and 6).

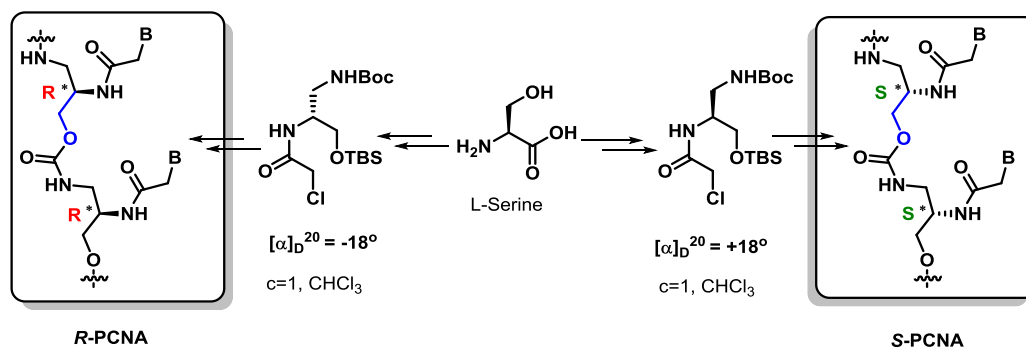
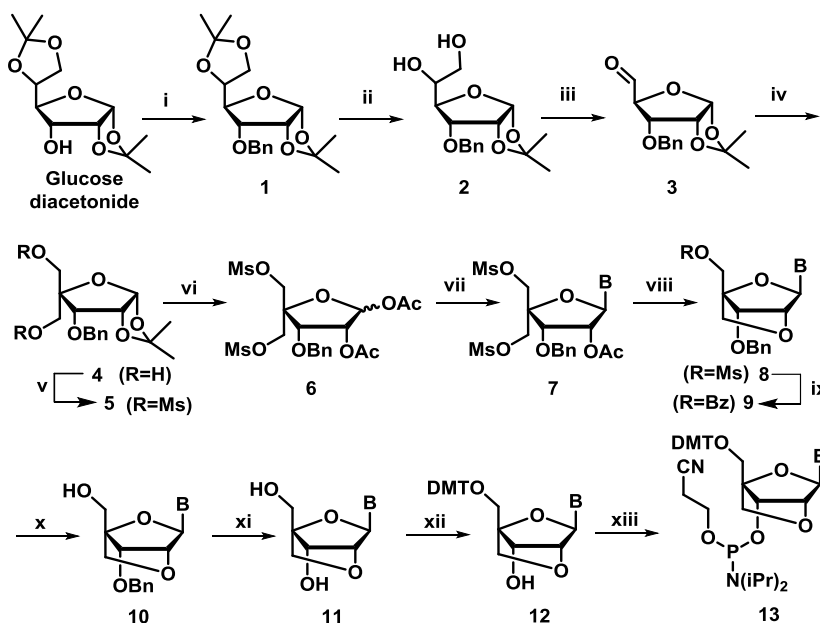


Figure 18 Synthetic outline showing key intermediates for the synthesis of R/S-PCNA from L-serine.

Scheme 4 Synthesis of conformationally locked thymidine phosphoramidite derivative



Reagents and conditions: (i) BnBr, NaH, dry DMF, quantitative yield (ii) 80% AcOH in water, quantitative yield (iii) NaIO₄, MeOH:H₂O (1:1) (iv) 2N NaOH, HCHO, dioxane, quantitative (v) MsCl, pyridine, 89% (vi) (a) 80% TFA; (b) Ac₂O, pyridine (vii) thymine, BSA, TMSOTf, MeCN (viii) 2N NaOH, dioxane:water (1:1) (ix) NaOBz, dry DMF (x) NaOH, THF (x) 20% Pd(OH)₂/C, HCO₂NH₄, MeOH (xii) DMT-Cl, pyridine (xiii) 2-cyanoethyl N,N-diisopropylchlorophosphoramidite, DIPEA, DCM.

Chapter 3

Table 3 Sequences synthesized for the study and their MALDI-TOF characterization

S. No.	Description	Sequence-Code (5'/N-3'/C)	Mass Calcd./Obsd.
1	DNA	ttttttt I-CF	2908.6/2907.2
2	PNA	ttttttt II-CF	2575.9/2579.6
3	R-PCNA	ttttttt III-CF	2759.9/2763.0
4	S-PCNA	ttttttt IV-CF	2759.9/2761.6
5	cDNA	gcaaaaaaacg V	3680.5/3676.7
6	non-cDNA	agtgacctac VI	3012.0/3018.4
7	DNA	cttcttcctt VII-CF	3457.0/3459.2
8	PNA	cttcttcctt VIII-CF	3106.4/3108.6
9	S-PCNA	cttcttcctt IX-CF	3267.9/3268.9
10	S-PCNA	atattattaatt X-CF	3936.2/3943.2
11	S-PCNA	actaaaactcat XI-CF	3900.5/3907.5
12	cDNA	aaggaagaag XII	3134.1/3140.1
13	cDNA	aattaataatat XIII	3651.5/3673.2
14	cDNA	atgagtttagt XIV	3690.5/3705.5
15	DNA	cctcttacctcagttaca XV-CF	5904.0/5904.4
16	Fully modified 2'- OMe RNA	ccucuuaccucaguuaca XVI-CF	6360.1/6360.9
17	DNA, $\underline{x} = t^{LNA}$	cc \underline{x} ct \underline{x} acc \underline{x} cagt \underline{x} aca XVII-CF	6016.0/6015.8
18	DNA, $\underline{x} = u^{S-AMP}$	cc \underline{x} ct \underline{x} acc \underline{x} cagt \underline{x} aca XVIII-CF	6260.2/6261.8
19	cDNA	tgtaactgaggtagagg XIX	5624.9/5623.9

3B.4 The study of short and different backboned-T₈-CF sequences towards probes development

3B.4.1 Studying the interactions of T₈-CF sequences with the GO using CD

We studied the binding abilities of the synthesized different backbone DNA-, PNA- and PCNA- octathymynyl (T₈) sequences with cDNA in buffer solutions using temperature-dependent UV-*T_m* experiments and found that PNA, R-PCNA and S-PCNA-T₈ show high *T_m* values indicating strong binding with cDNA (Table 4). We also confirmed the strong binding of these short R/S-PCNA sequences to DNA targets using gel retardation experiments (Chapter 3, Section A).

Table 4 Different backbone T₈ sequences and their UV-melting data

Description (8-mer)	Sequence-Code (5'/N-3'/C)	<i>T_m</i> (°C)* cDNA
DNA (phosphate)	ttttttt I-CF	~10
PNA (polyamide)	ttttttt II-CF	45.0
R-PCNA (carbamate)	ttttttt III-CF	50.1
S-PCNA (carbamate)	ttttttt IV-CF	56.2
cDNA	gcaaaaaaacg V	-
non-cDNA	agtgacctac VI	-

*UV-*T_m* values were measured by annealing 1 μM sequences with 1 μM cDNA in sodium phosphate buffer (0.01 M, pH 7.2) containing 10 mM NaCl, 0.1 mM EDTA and is an average of three independent experiments.

Binding of sequences **I-CF** to **IV-CF** with GO was followed by CD studies.^{28a} Perturbations in CD spectra were monitored to see the molecular interaction of individual sequences **I-CF** to **IV-CF** with GO (Figure 19). 1 μM solutions of **I-CF** to **IV-CF** in sodium phosphate buffer (0.01 M, pH 7.2) containing 10 mM NaCl were treated with 100 μg/mL of GO. The positive band at 280 nm for ss-DNA-T₈ **I-CF** was marginally enhanced upon addition of GO, similar to earlier reports (Figure 19a).^{28a} The achiral ss-PNA-T₈ **II-CF**, shows no characteristic CD, but upon addition of GO, we observed emergence of weak

CD signals near ~ 260 – 280 nm (Figure 19b). The changes in CD spectrum due to π - π /base-stacking interactions are reflected in this region of CD spectrum and these results do imply that these oligomers bind to GO in a definite manner.

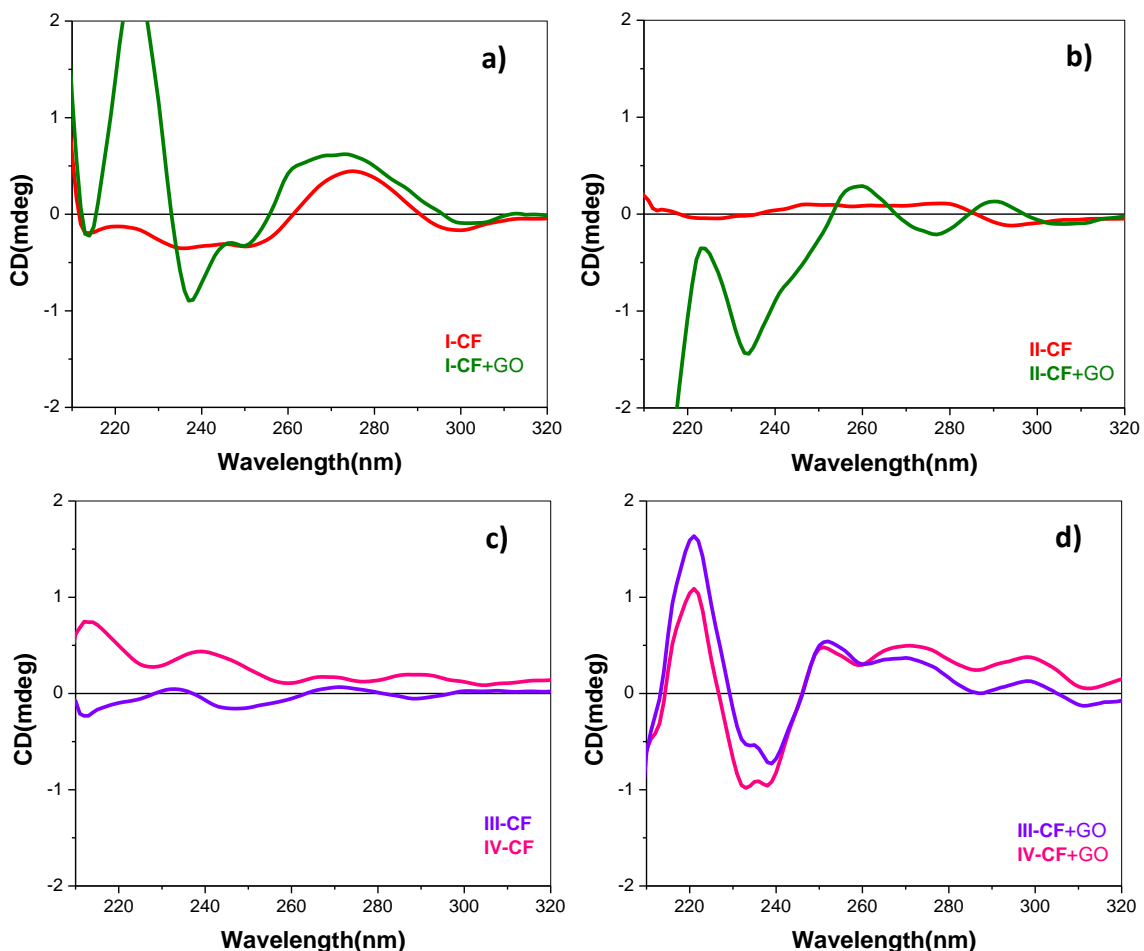


Figure 19 Perturbations in CD spectra of oligomers **I-CF** to **IV-CF** upon addition of GO (Buffer conditions: sodium phosphate buffer, 0.01 M, pH 7.2 containing 10 mM NaCl).

A mirror image pattern was observed for the CD spectra of enantiomeric ss-PCNA **III-CF** and **IV-CF** (Figure 19c). Surprisingly, upon addition of GO, the CD pattern of the two oligomers was almost identical near ~ 250 – 280 nm, indicating adaptation of very similar secondary structures for the *R/S*-PCNA oligomers when bound to GO (Figure 19d). It was quite interesting to observe that the backbone chirality of *R/S*-PCNA oligomers substantially affected the CD patterns upon adsorption on GO and even the achiral PNA

showed the emergence of a CD pattern in the base stacking region of the CD spectrum. The additional negative band at ~240 nm in all the spectra was also observed in earlier studies.^{28a} The CD studies thus provide evidence for the interaction of nucleic acid oligomers to GO surface.

3B.4.2 The fluorescence quenching and restoration of T₈-CF probes

We employed the CF-labelled oligomers **I-CF** to **IV-CF**, individually, to study the change in fluorescence emission intensity caused first by addition of GO, and later by challenging the GO complexes with the complementary ss-DNA target **V** under identical conditions. Thus, 50 nM solutions of **I-CF** to **IV-CF** were treated with 5 µg/mL of GO. The fluorescence of neutral backboned PNA- (**II-CF**) and *R/S*-PCNA- (**III-CF** and **IV-CF**) oligomers were quenched instantly (up to 90–95%, Figure 20b, 20c and 20d), whereas the fluorescence quenching of DNA **I-CF** was found to be much less (~ 20%, Figure 20a). Oligomer **I-CF** with its negatively-charged phosphodiester backbone interacted less competently with GO compared to the uncharged oligomers (**II-CF**, **III-CF** and **IV-CF**). This difference could be because of the high energy barrier while binding with GO due to electrostatic charge repulsions in the case of DNA. (An increased salt concentration was found to lead to an increase in the fluorescence quenching of **I-CF** with GO, Figure 21a) We then studied the restoration of fluorescence upon incremental addition of complementary ss-DNA **V** to estimate the desorption of each of the labelled probes from the GO surface (Figure 20).

There was almost no or very little fluorescence increase observed upon addition of **V** to **I-CF:GO**, **II-CF:GO** and **III-CF:GO** (~ 10%), whereas **IV-CF:GO** showed measurable (40%) restoration of the fluorescence intensity upon addition of 250 nM **V**, Figure 20d. The inability of sequences **I–III** to form duplexes on GO surface with cDNA could be one of the reasons for their failure to show increased fluorescence. Alternatively, the reason could be that the complexes of the T₈-CF sequences formed with cDNA also bind to the GO surface and therefore fluorescence cannot be restored. To substantiate either explanation, we then used the pre-formed complexes **I-CF:V**, **II-CF:V**, **III-CF:V** and **IV-**

CF:V in similar experiments (Figure 20). All the preformed duplexes show fluorescence intensity similar to the single strands in the absence of GO.

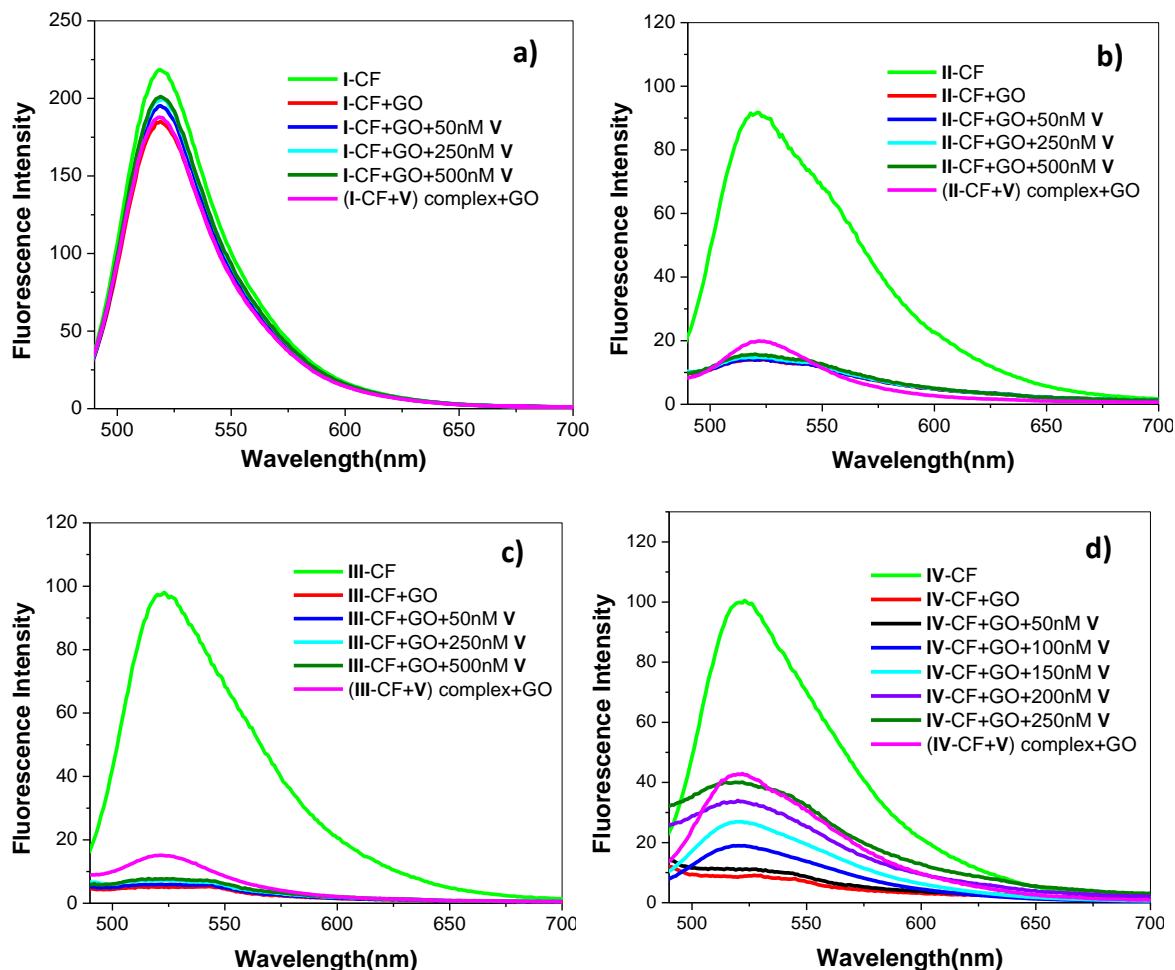


Figure 20 Fluorescence quenching of **I-CF** to **IV-CF** by GO and its restoration upon addition of cDNA **V**. (Buffer conditions: sodium phosphate buffer, 0.01 M, pH 7.2 containing 10 mM NaCl).

Upon adding GO, the fluorescence quenching of the preformed complexes was almost equal to that of ss- labelled oligomers in the case of complexes DNA **I-CF:V**, PNA **II-CF:V** and *R*-PCNA **III-CF:V** (Figure 20a, 20b and 20c). The quenching of complex fluorescence was about 60% in the case of **IV-CF:V** and the observed fluorescence intensity is equal to what was restored by adding the ss-cDNA **V** to the *S*-PCNA **IV-CF:GO** complex. These results indicate that the efficacy of fluorescence restoration

depends upon the difference in the binding of different complexes formed with cDNA **V** to the GO surface. The ds-DNA is known to bind to GO, albeit less than ss-DNA, and in this particular study, the short sequence of DNA backbone comparatively falls short for this difference to be expressed in terms of increased fluorescence. In the case of uncharged PNA, PNA:DNA complexes may have additional hydrogen bonding interactions through amides and π -stacking interactions with the oxidized graphene surface.⁵⁶ In the absence of electrostatic repulsions observed in DNA duplexes, PNA:DNA complexes were not desorbed efficiently from the GO surface and no signal could be observed. Such H-bonding interactions are expected to be less pronounced in carbamate oligomers as the H-bonded secondary structures in peptide oligomers are known to be disturbed by carbamate linkages.⁵⁷

3B.4.3 Specificity of the fluorescence restoration

As the initial fluorescence quenching of **I-CF** with GO was less with lower salt concentration (10 mM NaCl), we have examined whether an increased salt concentration (100 mM NaCl and 10 mM MgCl₂) would enhance the quenching efficiency as reported in earlier studies.³⁹ We therefore, repeated these studies at higher salt concentrations (Figure 21a). Upon increasing the salt concentration (Figure 21a and 21b), we found better adsorption of the labelled DNA when about 65% fluorescence was quenched by the added GO. Upon challenging with cDNA **V** we observed about 60–70% of immediate fluorescence restoration. This could be due to improved DNA duplex stability (**I-CF:V**) at higher salt concentration, in addition to relatively less adsorption of duplex DNA on the GO surface³⁹ (Figure 15, Path A). Increased salt concentration did not have appreciable effect on the fluorescence quenching and restoration of the neutral backbone probes (**II-CF**, **III-CF** and **IV-CF**). Among fluorescence quenched complexes of **I-CF:GO**, **II-CF:GO**, **III-CF:GO** and **IV-CF:GO**, addition of cDNA **V** could restore the prominent fluorescence only in case of **I-CF** and **IV-CF** probes.

To substantiate this reasoning, we treated the **I-CF:GO** and **IV-CF:GO** complexes with a non-cDNA **VI** in a similar fashion. As was observed in earlier reported

experiments,^{37, 39} fluorescence restoration was observed in case of **I-CF:GO** even with the addition of non-cDNA **VI**, Figure 21b. This could be due to a simple reversible adsorption process for the added non-cDNA **VI** sequence *vis-a-vis* **I-CF** (Figure 15, Path B) and would lead to a nonspecific fluorescence signal.³⁹ On the other hand, there was no fluorescence restoration observed in case of **IV-CF:GO** with the addition of non-cDNA **VI** (Figure 21c). This observation together with the prominent fluorescence restoration observed with the cDNA **V** addition suggests that path B of Figure 15 does not operate in case of probe **IV-CF**.

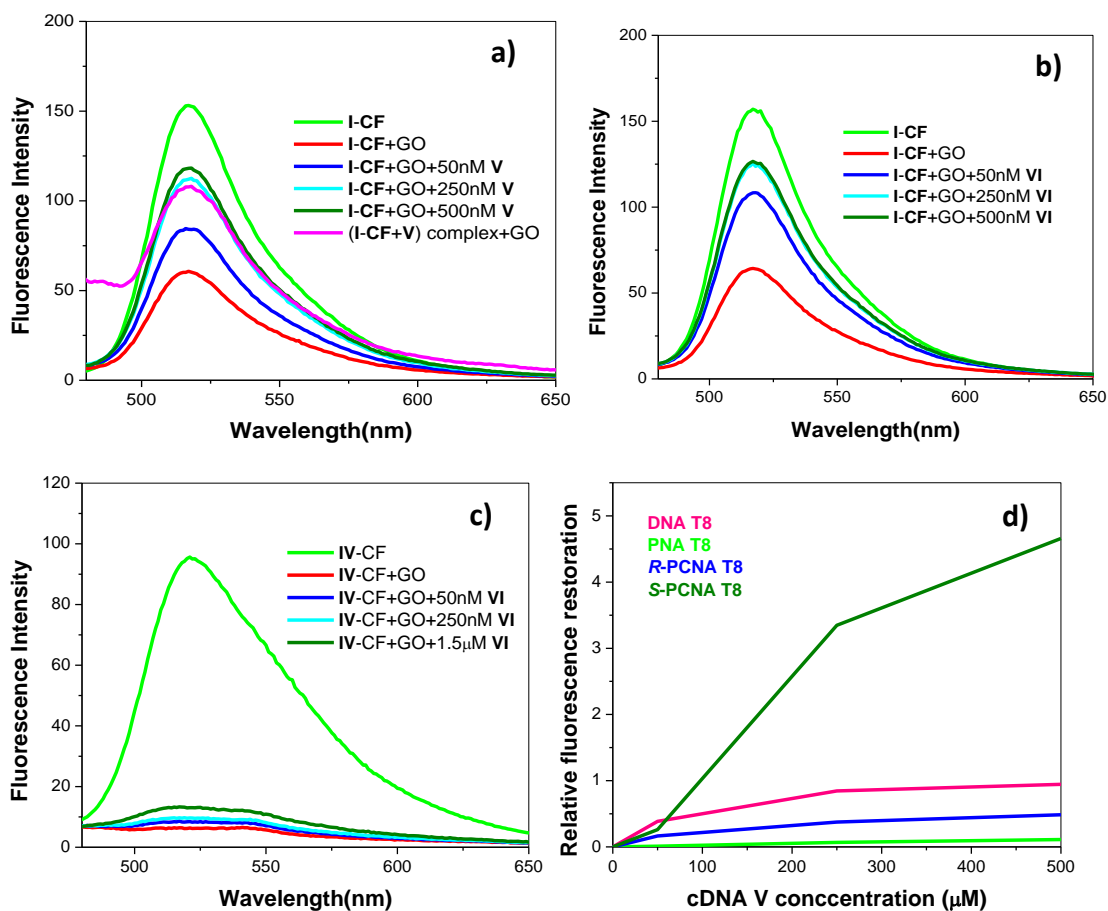


Figure 21 Fluorescence quenching of **I-CF** by GO and its restoration upon addition of **a)** cDNA **V** **b)** non-cDNA **VI**. [Buffer conditions: sodium phosphate buffer (0.01M, pH 7.2) containing 100 mM NaCl and 10 mM MgCl₂] **c)** no prominent fluorescence restoration of **IV-CF** with the addition of non-cDNA **VI** and **d)** comparative fluorescence restoration of the T₈-CF probes.

The reason for the differential interaction of GO with **III-CF:V** and **IV-CF:V** could be the different kinetics of binding and desorption of the duplexes from GO, arising from the different chirality of the PCNA oligomers and structural differences between their duplexes with DNA. However the CD spectral studies were not indicative of the differences for **III-CF:V** and **IV-CF:V** duplexes, as CD spectra of either complex were found to be almost identical, and were dominated by CD signals due to chirality of DNA. From these preliminary results we have demonstrated that a modified DNA backbone such as *R*-PCNA which is chiral and uncharged is better as a probe for DNA detection than natural DNA (induces non-specific fluorescence restoration besides nuclease-susceptibility) or PNA.

3B.5 Studies extended to mixed pyrimidine/purine oligomeric probes

Additional experiments with respect to oligomers of modified backbones, their length and sequence were carried out to reveal the generality and applicability of the above result in developing superior capture probes. We have synthesized the three labelled sequences of 10-mer in length (N/5'-CTTCTTCCTT-3'/C) with the backbones of phosphate- (**VII-CF**), amide/polyamide- (**VIII-CF**) and the carbamate *S*-PCNA- (**IX-CF**). UV- T_m studies were also performed to determine the duplex stability with the cDNA (**XII**) (Table, 5).

The fluorescence quenching and restoration of **VII-CF** and **VIII-CF** from the GO surface, upon challenging with the cDNA (**XII**) was similar to that observed for the T₈-CF sequences (**I-CF** and **II-CF**), i.e. the non-specific restoration of fluorescence in case of phosphate backboneed-**VII-CF** and no fluorescence restoration in case of amide backboneed-**VIII-CF**. Surprisingly, in contrast to the *S*-PCNA-T₈-CF (**IV-CF**) which specifically restored the fluorescence with cDNA **V** addition, the same chiral carbamate backbone containing **IX-CF** couldn't be desorbed from GO surface upon challenging **IX-CF:GO** complex with the cDNA (**XII**), resulting in no fluorescence restoration.

Table 5 Modified oligomers synthesized for the extension studies and their UV- T_m data

Description (length)	Sequence-Code (5'/N-3'/C)	T_m °C* cDNA
DNA (10)	cttcttcctt VII-CF	29.1
PNA (10)	cttcttcctt VIII-CF	52.0
S-PCNA (10)	cttcttcctt IX-CF	57.6
S-PCNA (12)	atattattaatt X-CF	56.8
S-PCNA (12)	actaaaactcat XI-CF	59.2
cDNA	aaggaagaag XII	-
cDNA	aattaataatat XIII	-
cDNA	atgagtttagt XIV	-

*UV- T_m values were measured by annealing 1 μ M sequences with 1 μ M cDNA in sodium phosphate buffer (0.01M, pH 7.2) containing 10mM NaCl, 0.1mM EDTA and is an average of three independent experiments

As homo pyrimidine sequences **VII-CF** – **VIII-CF** contains four cytosine residues, with the possibility of cytosine amines getting protonated we thought of repeating the studies at pH of 5.8, 6.6 and 7.4, to test the contribution of cytosine protonation to the observed result. Repetition of the studies at different pH also did not deliver the cDNA induced fluorescence restoration. In addition, we have synthesized two other 12-mer sequences to repeat the similar studies, containing the same backbone as that of S-PCNA-**T₈-CF**, **XI-CF** (N-ATA TTA TTA ATT-C) and **XII-CF** (N-ACT AAA ACT CAT-C). Though both the sequences formed stable duplexes with cDNA, both the sequences disappointed us in the fluorescence restoration experiments. Both the neutral backboneed **X-CF** and **XI-CF** were adsorbed efficiently on to the GO surface even with 3 μ g/mL of GO (Figure 22). Further, upon challenging with the 50nM cDNA there was little fluorescence restoration in case of **X-CF** and no restoration in case of **XI-CF**. Unfortunately there was no incremental/prominent restoration with the successive addition of cDNA.

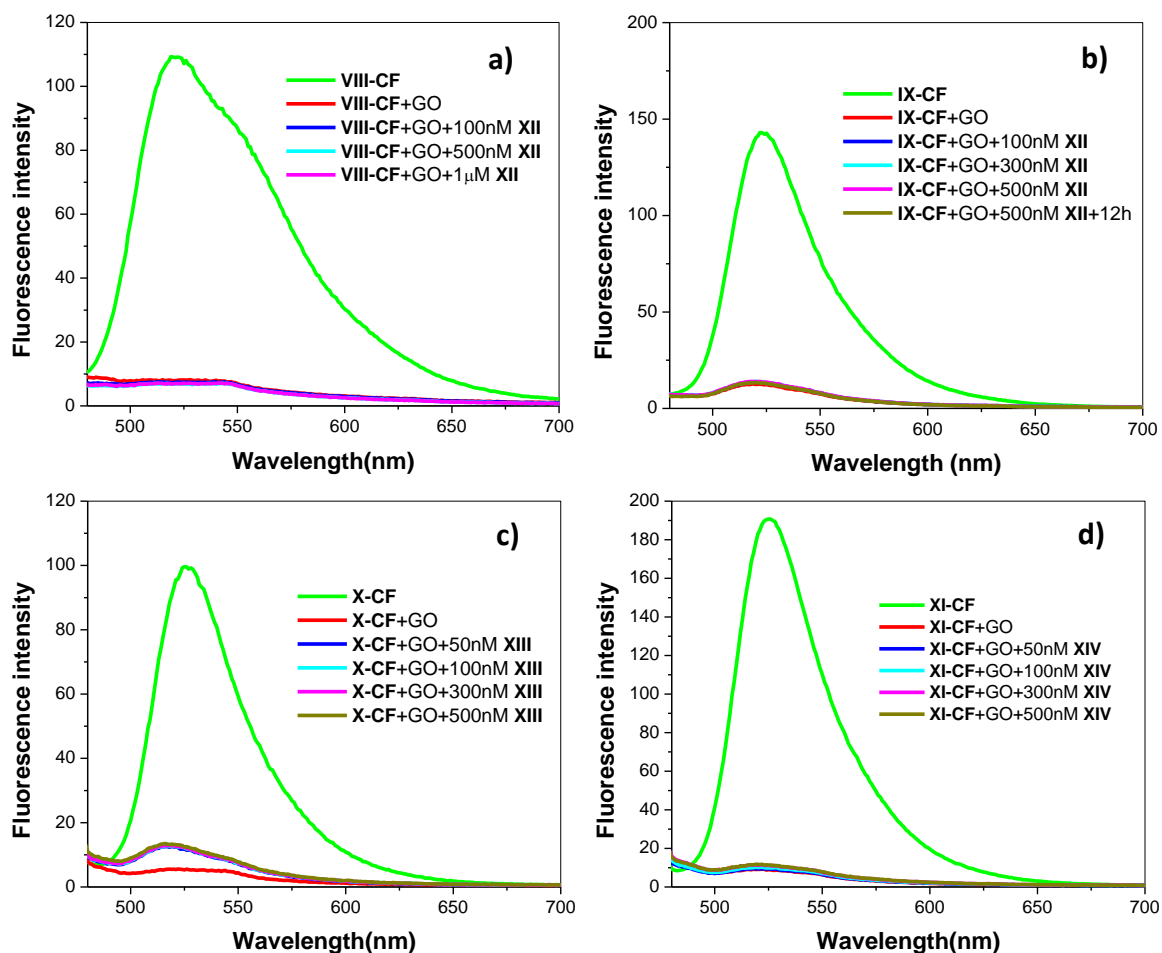


Figure 22 Fluorescence quenching of VIII-CF to XI-CF by GO and its restoration upon addition of cDNA. (Buffer conditions: sodium phosphate buffer, 0.01 M, pH 7.2 containing 10 mM NaCl)

3B.6 Thermal desorption studies of the probes from GO surface

In addition to the cDNA-induced desorption discussed so far, we envisioned that the temperature dependent desorption studies would give further insights into our current understanding of the interactions of ss/ds DNA with GO. Therefore, the thermal dissociation of adsorbed labeled probes I-CF to IV-CF and VII-CF to XI-CF were carried out. In a typical experiment, fluorescence of the labeled probe was quenched with the adequate amount of GO (for phosphate backbone $\sim 100 \mu\text{g}/\text{mL}$ GO and 3-5 $\mu\text{g}/\text{mL}$ GO for the neutral backbones) and the fluorescence was recorded at $\lambda_{\text{em max}}$ at every 2.0 $^{\circ}\text{C}$ rise in temperature.

As expected the thermal dissociation of the probes led to increased fluorescence, which followed a sigmoidal pattern (Figure 23). To find out the correlation between the cDNA-induced desorption and temperature-induced desorption, we recorded the fluorescence in presence of cDNA, in a similar way as described above. The temperature-dependent fluorescence changes of the four T₈-CF sequences (**I-CF** to **IV-CF**) are shown in Figure 23a and their first order derivatives are given in Figure 23b. The temperature-dependent fluorescence changes of the remaining probes which were synthesized for the extension studies (**VII-CF** to **XI-CF**) are shown in Figure 23c and their first order derivatives are as shown in Figure 23d.

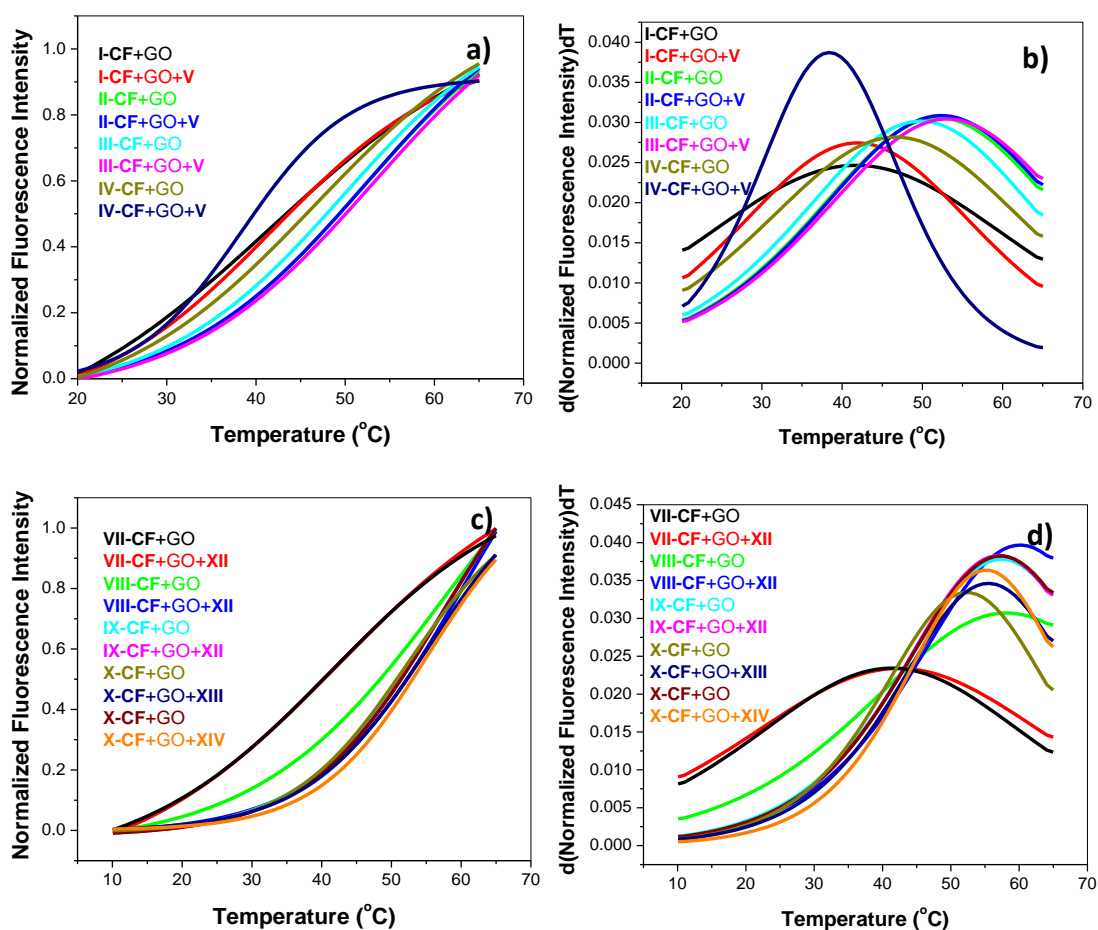


Figure 23 Thermal desorption of the labeled probes (**I-CF** to **IV-CF** and **VII-CF** to **XI-CF**) from the GO surface in presence and absence of cDNA (V/ XII/ XIII/ XIV) following sigmoidal plots (a and c) and their first order derivatives (b and d)

The maxima of the first order derivatives of change in the fluorescence intensity during thermal-induced desorption of the labeled probes (**I-CF** to **IV-CF** and **VII-CF** to **XI-CF**) from GO are presented in Table 6. Thermal desorption studies revealed that all the synthesized probes **I-CF** to **XI-CF** desorbed from the GO surface approximately at the same temperature independent of presence of cDNA, except **IV-CF**. Only **IV-CF** has significant difference in desorption temperature (-8.2 °C) in the presence and absence of cDNA (38.4 °C and 46.6 °C respectively). This result suggests the difference in binding strength of ss-**IV-CF** and its duplex with cDNA while binding to GO surface. In all the other duplexes, the temperature at which fluorescence restoration occurs upon heating is almost equal for ss probe and the duplex with cDNA. Our experimental results with DNA and probes with different backbones such as PNA, *R*-PCNA and *S*-PCNA therefore may be explained based on these results. (Figures 20, 21 and 22).

Table 6 Maxima of the first order derivatives of the change in fluorescence during thermal-induced desorption of the labeled probes from GO

No.	Probe (type of backbone, length)	Maxima of dF/dT	
		Probe+GO	Probe+GO+cDNA
1	I-CF (phosphate, 8-mer)	41.7	41.8
2	II-CF (peptide, 8-mer)	51.9	52.3
3	III-CF (<i>R</i> -carbamate, 8-mer)	49.8	53.0
4	IV-CF (<i>S</i> -carbamate, 8-mer)	46.6	38.4
5	VII-CF (phosphate, 10-mer)	42.4	41.5
6	VIII-CF (peptide, 10-mer)	58.1	60.1
7	IX-CF (<i>S</i> -carbamate, 10-mer)	55.8	56.7
8	X-CF (<i>S</i> -carbamate, 12-mer)	52.3	55.4
9	XI-CF (<i>S</i> -carbamate, 12-mer)	53.0	55.1

3B.7 Evaluation of longer DNA/DNA modified probes

As the short DNA sequence probes bind strongly to the GO surface,³⁷ we conceived to synthesize longer modified DNA probes which are capable of either forming strong

duplexes with the target cDNA or exerting additional electronic effects that might help in the desorption of probes from GO. Four labeled sequences each carrying no/different sites of modifications were synthesized, purified and characterized by MALDI-TOF. Each probe has different binding abilities to cDNA, studied by UV- T_m experiments (Table 7). The duplex stability order of the studied probes with cDNA is as follows: **XVII-CF** (4-units of T^{LNA}) > **XV-CF** (unmodified DNA) > **XVI-CF** (fully 2'-OMe modified) > **XVIII-CF** (4-units of U^{AMP}). In this case, the obtained duplex stabilities could not be correlated to the cDNA-induced desorption from GO surface. The probes **XVI-CF** and **XVIII-CF**, though displaying the lower duplex stability to cDNA than **XV-CF** (unmodified DNA) they are anticipated to have additional electronic effects which would help in cDNA-induced desorption. The initial fluorescence quenching of all probes is efficient with the addition of 10 $\mu\text{g/mL}$ of GO, **XVI-CF**, **XVII-CF** and **XVIII-CF** being slightly better than **XV-CF**. The restoration of fluorescence with the incremental addition of cDNA is studied for all the probes under identical conditions.

Table 7 Phosphate backbone modified oligomers synthesized for the extension studies and their UV- T_m data

Description (18 mer)	Sequence-Code (5'/N-3'/C)	T_m °C* cDNA
DNA	cctcttacctcagttaca XV-CF	54.8
Fully modified 2'-OMe RNA	ccucuuaccucaguuaca XVI-CF	53.4
DNA, $\underline{x} = t^{LNA}$	cc \underline{x} ct \underline{x} acc \underline{x} cagt \underline{x} aca XVII-CF	64.1
DNA, $\underline{x} = u^{S-AMP}$	cc \underline{x} ct \underline{x} acc \underline{x} cagt \underline{x} aca XVIII-CF	46.1
cDNA	tgtaactgaggtgaagagg XIX	-

*UV- T_m values were measured by annealing 1 μM sequences with 1 μM cDNA in sodium phosphate buffer (0.01M, pH 7.2, 100mM NaCl, 0.1mM EDTA) and is an average of three experiments

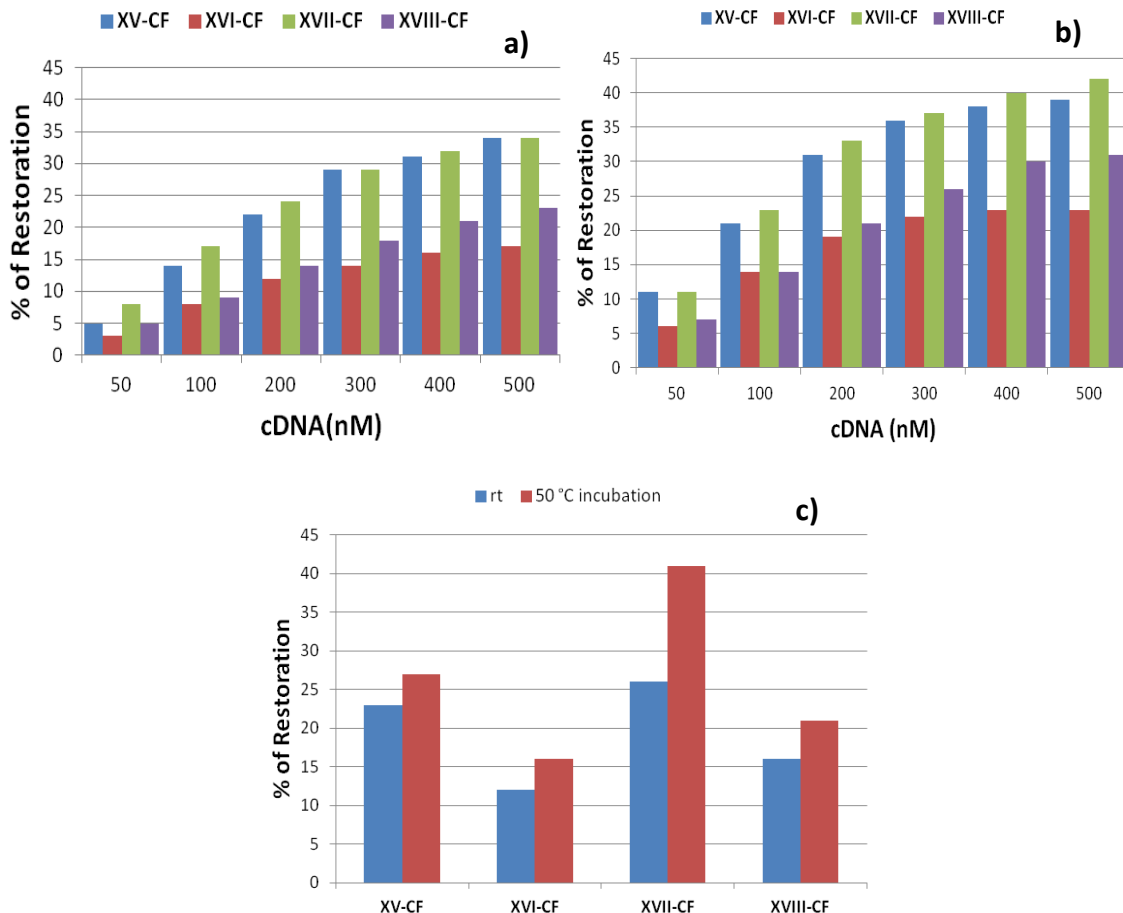


Figure 24 % Fluorescence restoration with the successive addition of cDNA **XIX** in regular (a) and high salt (b) buffer conditions and also the relative % restoration at room temperature and 50 °C incubation for 30 minutes [Buffer conditions: regular salt- sodium phosphate buffer (0.01 M, pH 7.2) containing 100 mM NaCl and 10 mM MgCl₂; high salt - sodium phosphate buffer (0.01M, pH 7.2) containing 1000 mM NaCl]

Though the fluorescence restoration upon challenging with cDNA **XIX** was not prominent, the trend was dominated by the **XVII-CF** probe (Figure 24), which could be easily attributed to high duplex stability imparted by the LNA-modifications in this probe (Table 5, entry 3). The increasing salt concentration and the incubation at 50 °C for 30 minutes, also did not result in appreciable change in the fluorescence restoration (Figure 24). Thus, the preliminary cDNA induced fluorescence restoration of these probes indicates that high duplex binding and higher duplex desorption from GO are the key contributing

parameters in improving the detection. It appears that modulating the physical parameters like duplex strength, temperature, buffer conditions, etc. might be beneficial in developing the probes. These studies are in progress in our laboratory.

3B.8 Conclusion

- ✓ For the first time we have introduced the use of synthetic analogues in conjunction with GO to develop the nucleic acid-based capture probes.
- ✓ Initially we have demonstrated the interactions of short-, labeled nucleic acids (T₈-CF) of varying backbones like phosphate, poly amide and carbamate with GO and shown that the neutral backboneed nucleic acids were adsorbed to GO very efficiently.
- ✓ Subsequently, cDNA-induced specific desorption of the probes from the GO surface was performed. Among the varied backboneed T₈-CF sequences, only the *S*-PCNA-T₈-CF (**IV-CF**) showed the selective prominent and significant fluorescence restoration upon the cDNA (**V**) addition.
- ✓ The studies were extended to the mixed sequences of 10-mer (containing phosphate, peptide and carbamate backbones) and 12-mer (having the *S*-PCNA backbone). Unfortunately, these sequences did not show any fluorescence restoration upon the addition of cDNA.
- ✓ In case of 10-mer mixed homopyrimidine sequences (**VII-CF**, **VIII-CF** and **IX-CF**), change of pH also did not have any effect on the cDNA-induced desorption of the probes.
- ✓ Thermal desorption of the synthesized short sequence probes from the GO surface were studied in the presence and absence of cDNA. The observed results are in well accordance with the cDNA-induced spontaneous desorption of the probes from GO surface at room temperature.
- ✓ We synthesized new DNA probes containing modifications of our design and LNA. However the fluorescence restoration studies need to be optimized.
- ✓ In conclusion, our studies illustrate the possibility of using novel, stable synthetic DNA analogues and will help in better understanding of the nucleic acids interaction with GO,^{45,58} stimulating further research in this important area.

3B.9 Experimental Section

General

PCNA and PNA sequences were synthesized manually using solid phase peptide synthesis and the DNA sequences were made using a Bioautomation Mermade 4 DNA synthesizer. All the oligomers were purified by RP-HPLC on a C18 column and characterized by MALDI-TOF spectrometry. MALDI-TOF spectra were obtained on a Voyager-De-STR (Applied Biosystems). CHCA (α -Cyano-4-hydroxycinnamic acid) matrix was used to analyze PCNA/PNA samples and THAP (2,4,6-trihydroxy acetophenone)/ammonium citrate were used to analyze the DNA samples. UV absorbance and melting studies were performed on a Varian Cary 300 UV-VIS spectrophotometer. Circular Dichroism (CD) analysis was performed on a JASCO J-715 spectrophotometer using a cell of 10mm path length. Spectra were recorded as accumulations of 5 scans using a scan speed of 200 nm/min, resolution of 1.0 nm, band-width 1.0 nm and a response of 1 sec. Spectra were smoothed and plotted using Origin Pro 6.1. X-ray powder diffraction studies were performed using $\text{CuK}\alpha$ ($\lambda = 1.54 \text{ \AA}$) radiation on a Philips PW1830 instrument operating at 40 kV and a current of 30 mA at room temperature. Diffraction patterns were collected at a step of 0.020 (2θ) after background subtraction with the help of a linear interpolation method. The samples were prepared as thin films on glass substrates. All Raman spectroscopy measurements were carried out at room temperature on an HR 800 Raman spectrophotometer (Jobin Yvon Horiba, France) using monochromatic radiation emitted by a He-Ne laser (632.8 nm) operating at 20 mW. The experiment was repeated three times to verify the consistency of the spectra. The samples were prepared as thin films on glass substrates. Fluorescence experiments were performed on Varian's Cary Eclipse Fluorescence Spectrophotometer. Fluorescence emission intensity was recorded from 480 to 700nm, with a maximum observed at 520nm, using the excitation wavelength of 470nm. CD and fluorescence experiments were performed in 10mM sodium phosphate buffer containing 10mM NaCl at pH=7.2. The fluorescence of the complexes was recorded at $\sim 6^\circ\text{C}$.

Synthesis and characterization of graphene oxide

Graphene oxide (GO) was prepared from graphite powder according to modified Hummers' methods.^{26, 28c, 37} Accordingly, graphite powder (1g) was put into sulfuric acid (25 mL). KMnO_4 (3 g) was added at 0 °C gingerly under stirring. After addition of KMnO_4 , the ice-bath was removed and the suspension was stirred at 35 °C for 30 min. Deionized water (46 mL) was added slowly, and the temperature of the suspension was increased to 98 °C and kept for 15 min, then the suspension was further diluted to 140mL with warm deionized water. Upon treating with 30% H_2O_2 (1 mL), the colour of the suspension changed from black to yellow. The yellow suspension was separated by centrifugation and washed repeatedly with HCl: water (1: 10). The final product was dried over phosphorus pentoxide *in vacuo* for two days. The material was characterized by XRD, Raman, CD and UV spectroscopy.

UV- T_m studies of synthetic oligomeric probes with cDNA

The complexation of PCNA **III-CF**, **IV-CF**, **IX-CF**, **X-CF** and **XI-CF** with their respective complementary DNA was studied by UV melting. The PCNA oligomer and cDNA of 1 μM concentration each were mixed together in 10mM sodium phosphate buffer (pH=7.2) containing 10mM NaCl. The samples were annealed by heating at 90 °C for 2 min, followed by slow cooling to room temperature and refrigeration at 4°C overnight prior to the experiment. The T_m values obtained are an average of three independent measurements and are accurate to $\pm 1^\circ\text{C}$. The UV absorption at 260nm was monitored as a function of temperature, when sigmoid transitions were observed. The T_m values were obtained from the peaks of the first derivative plots (**Figure 25**).

The complexation of **XV-CF**, **XVI-CF**, **XVII-CF** and **XVIII-CF** with complementary DNA were studied by UV melting, under the similar conditions as described for PCNA oligomers, the only difference being salt concentration used (100 mM NaCl instead of 10 mM NaCl), **Figure 26**.

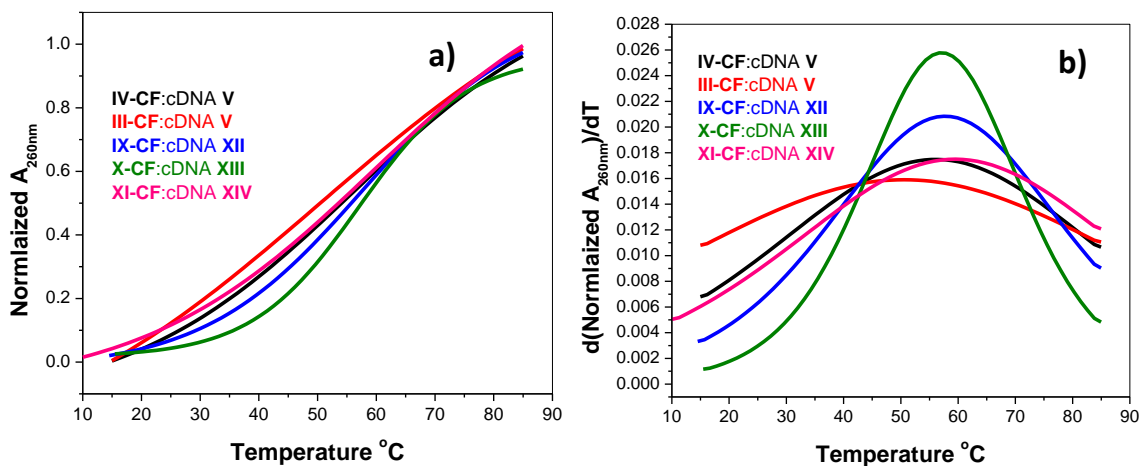


Figure 25 UV melting profiles (a) and first order derivative plots (b) of S-PCNA probes with cDNA

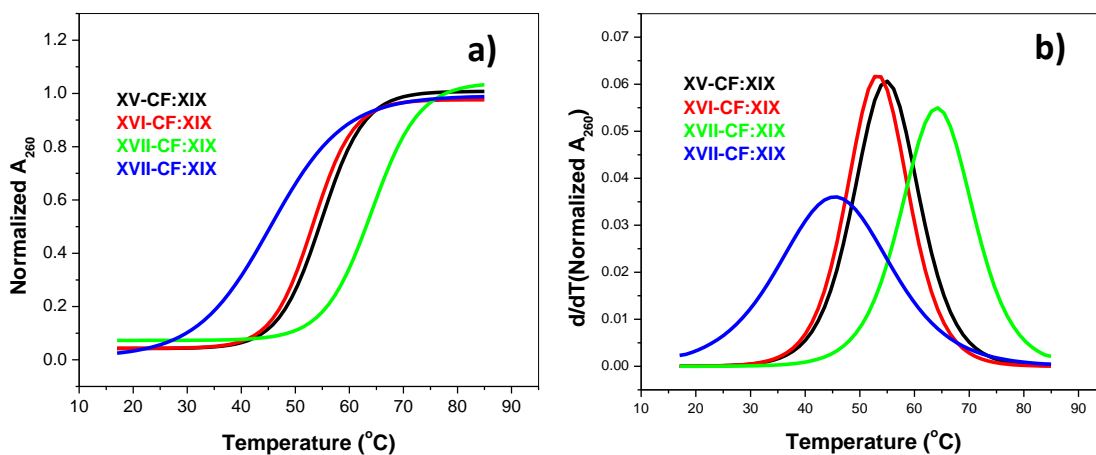
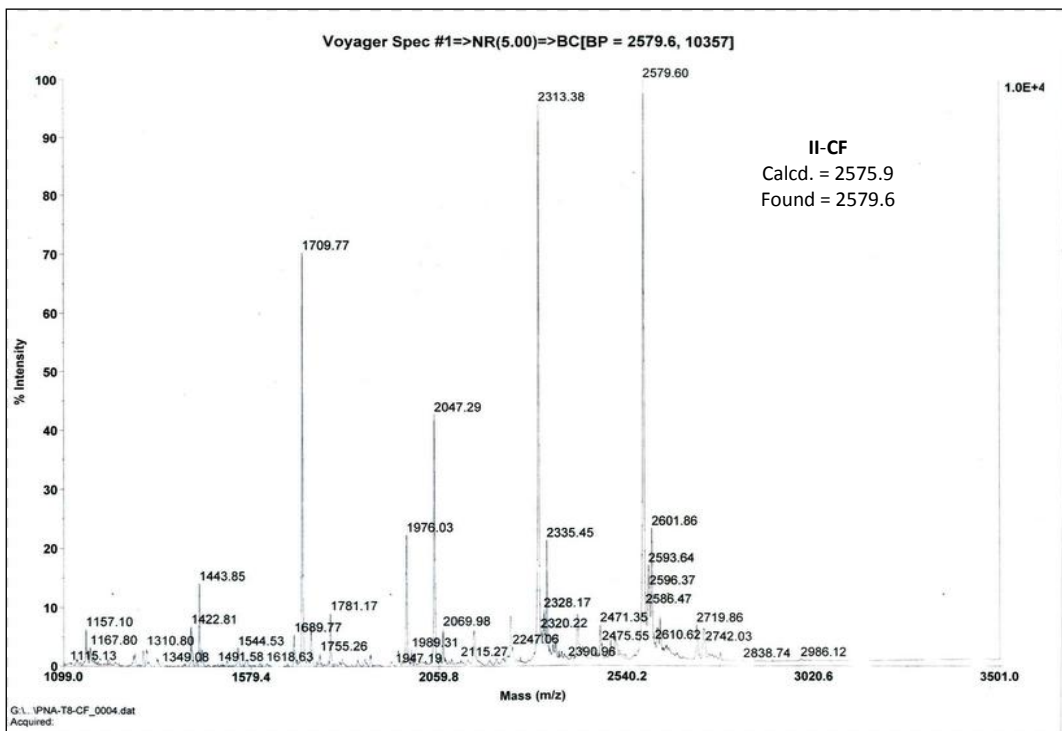
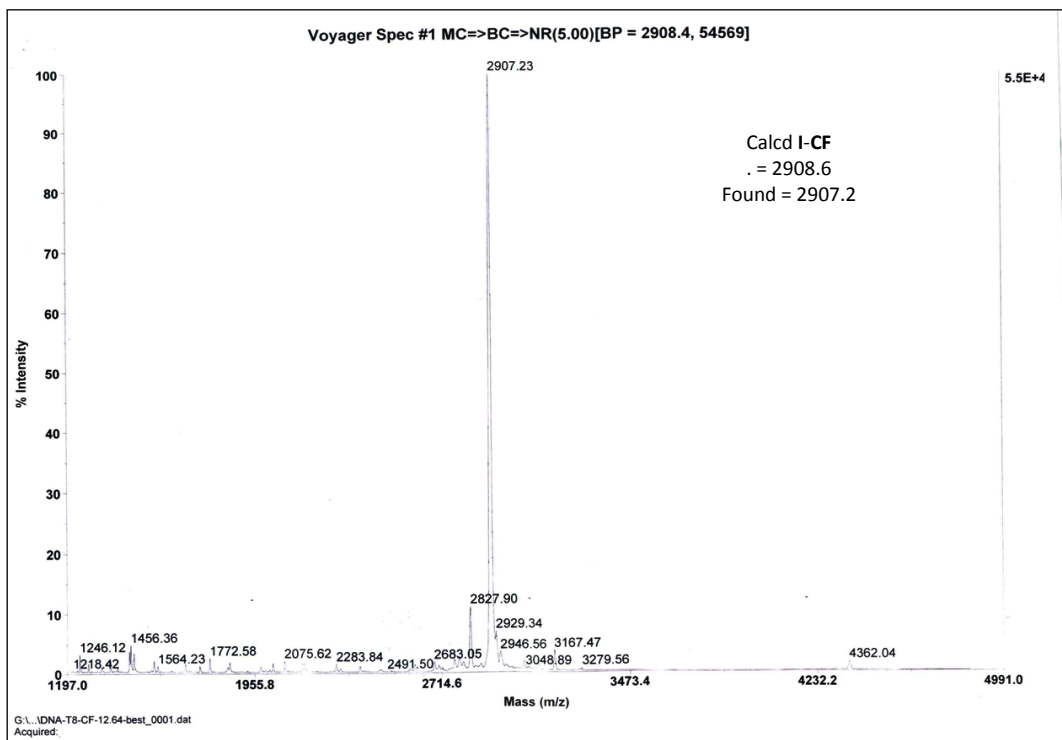


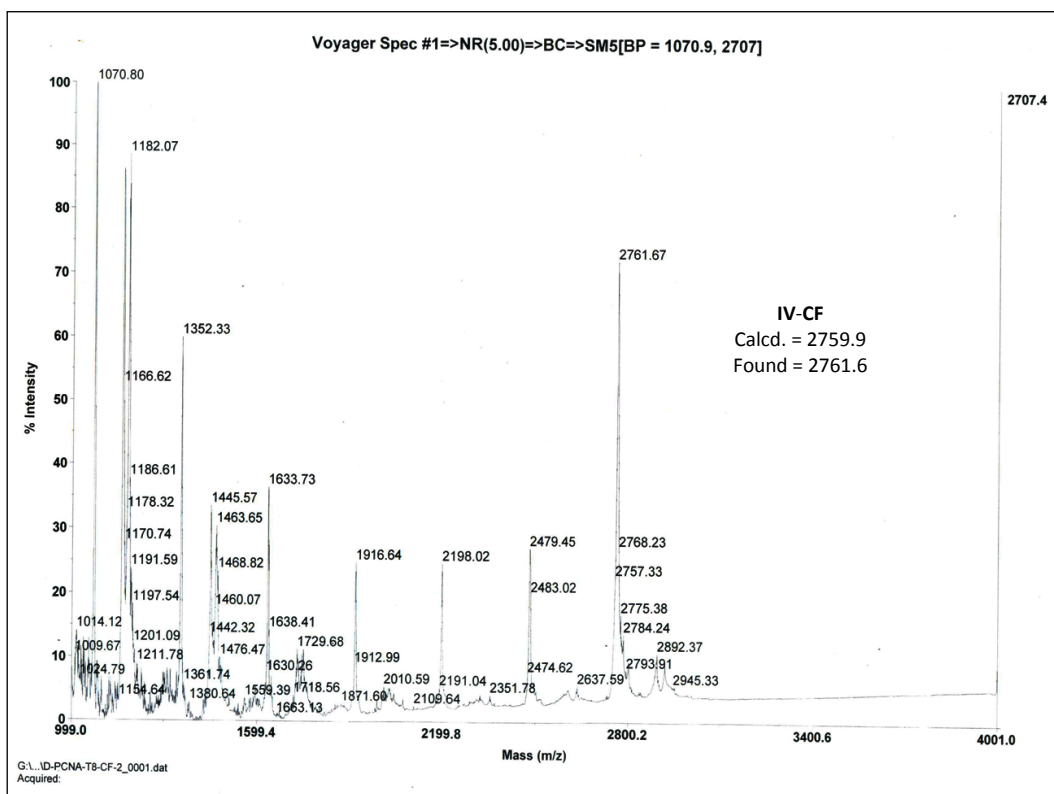
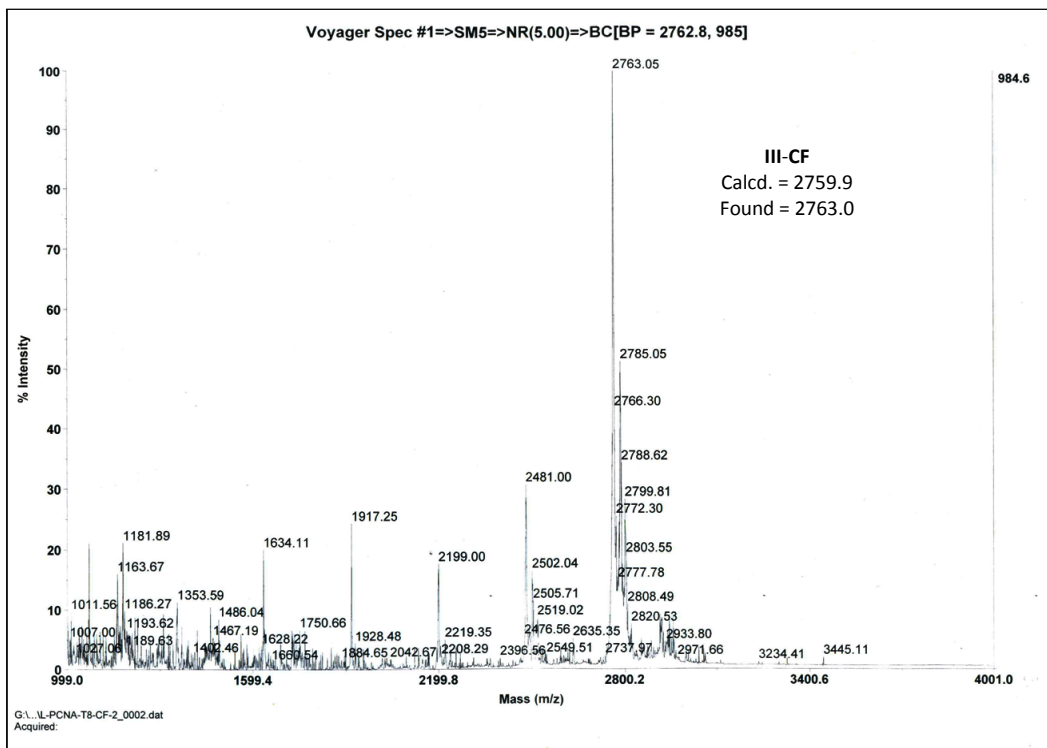
Figure 26 UV melting profiles (a) and first order derivative plots (b) of DNA/-analogue probes with cDNA

3B.10 Appendix

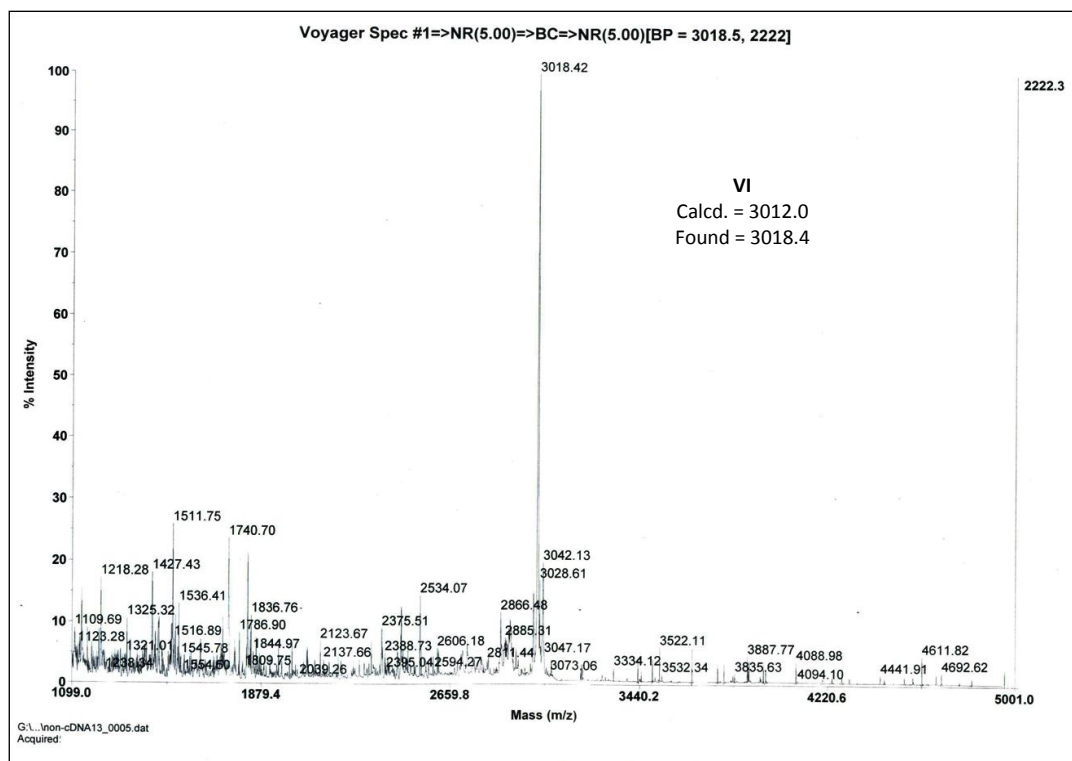
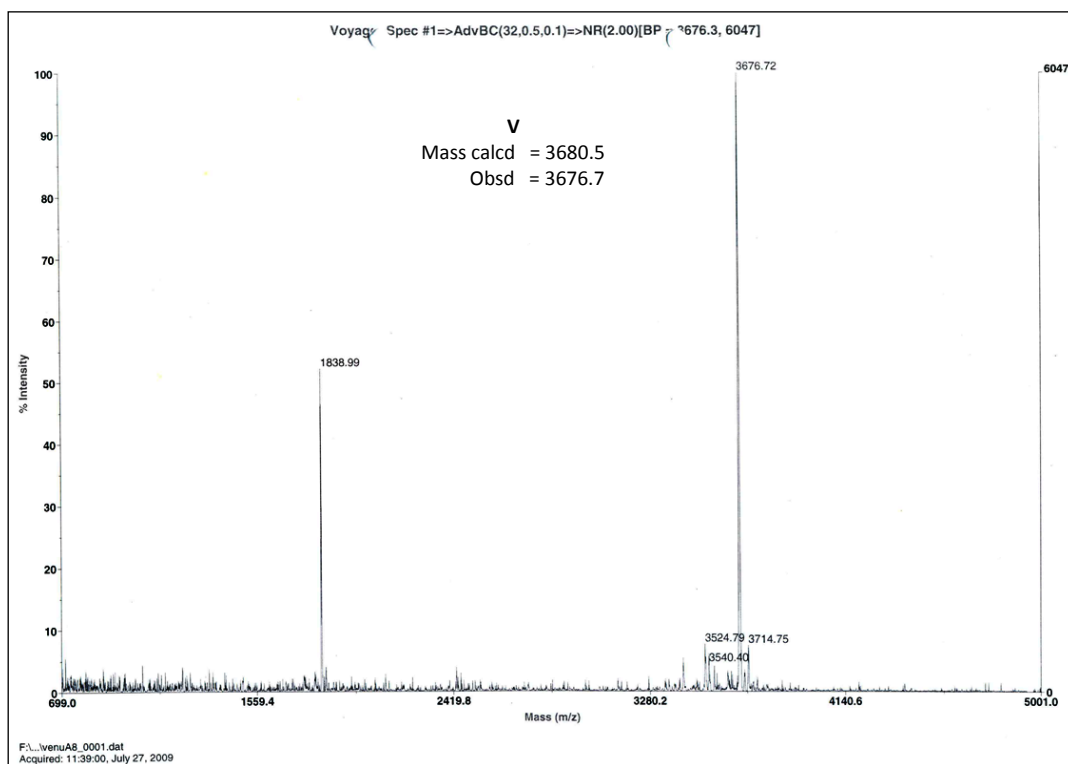
Sequences - MALDI-TOF Spectra	Page No.
I-CF & II-CF	249
III-CF & IV-CF	250
V & VI-CF	251
VIII-CF & IX -CF	252
X-CF & XI-CF	253
XII & XIII	254
XIV & XV-CF	255
XVI-CF & XVII-CF	256
XVIII-CF & XIX	257



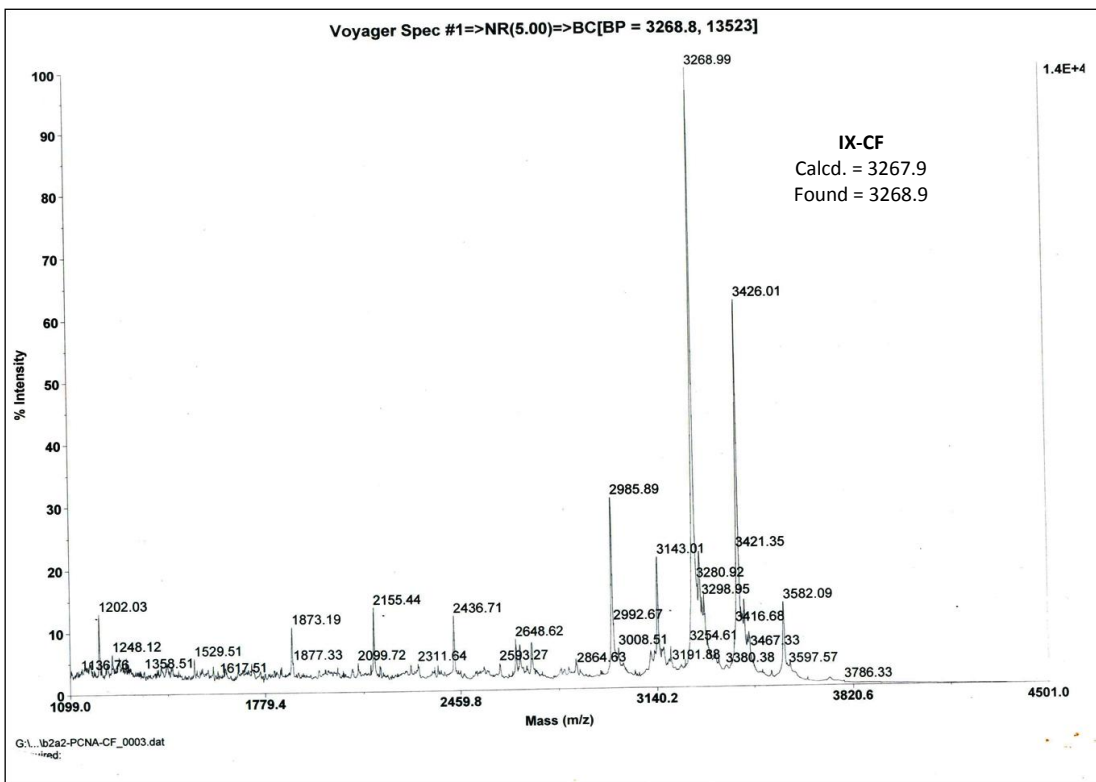
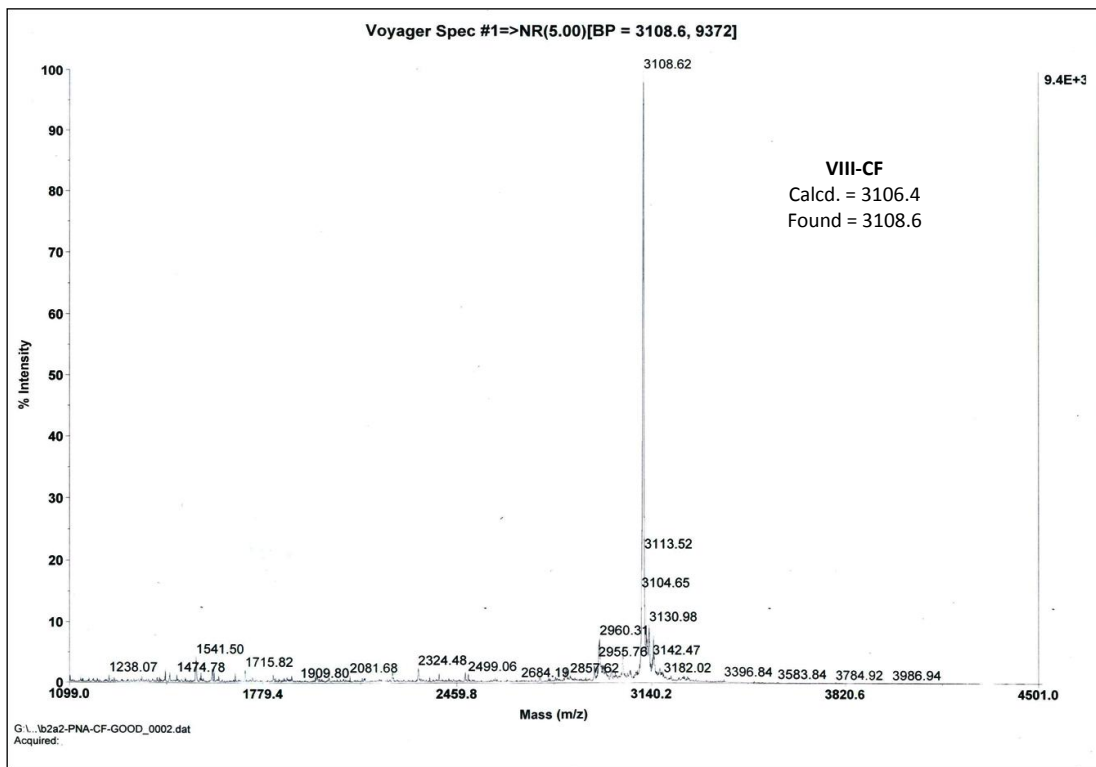
Chapter 3



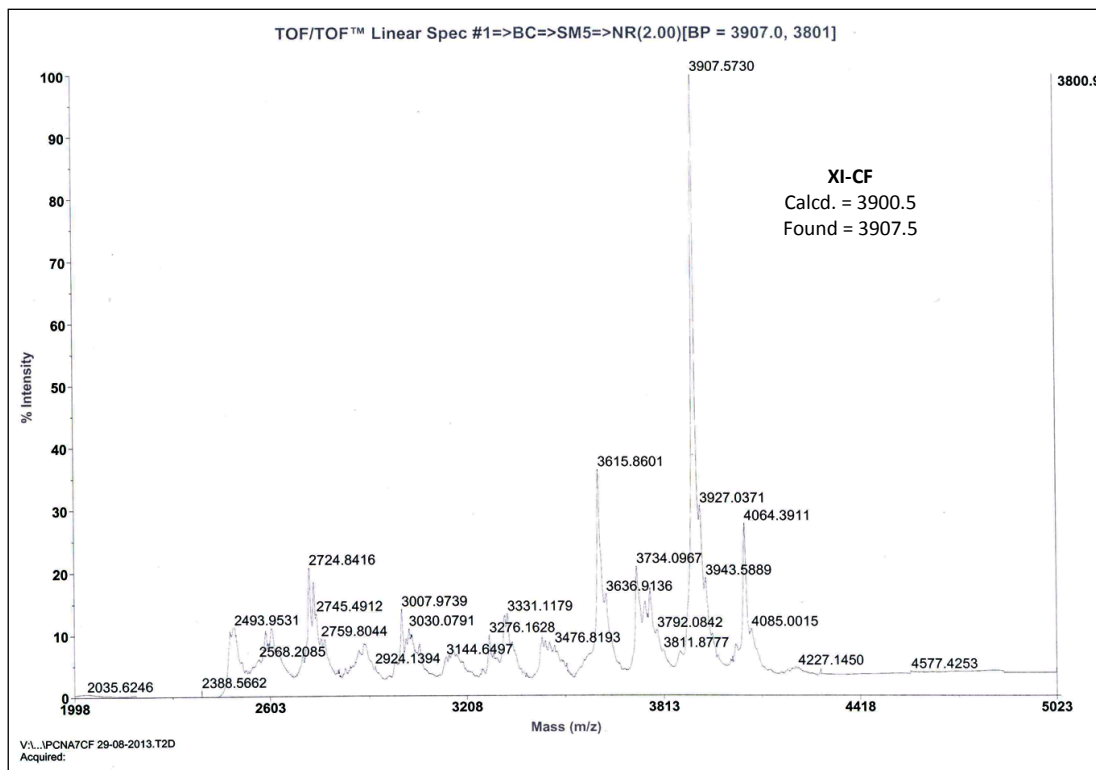
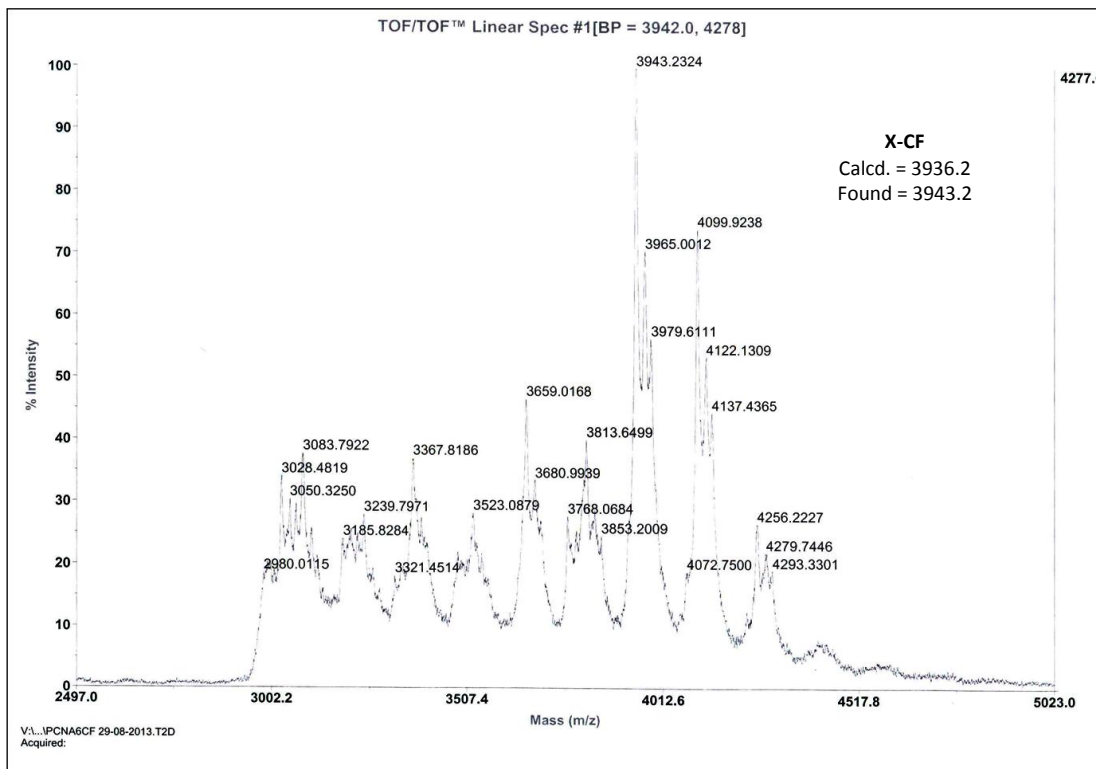
Chapter 3



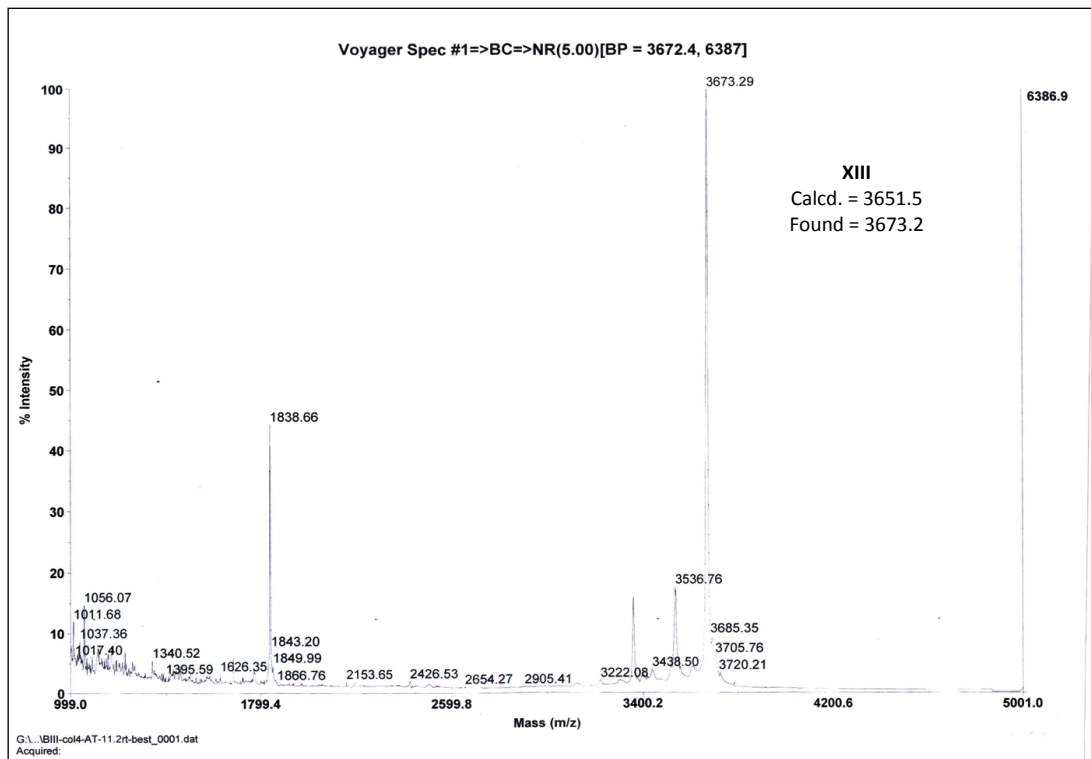
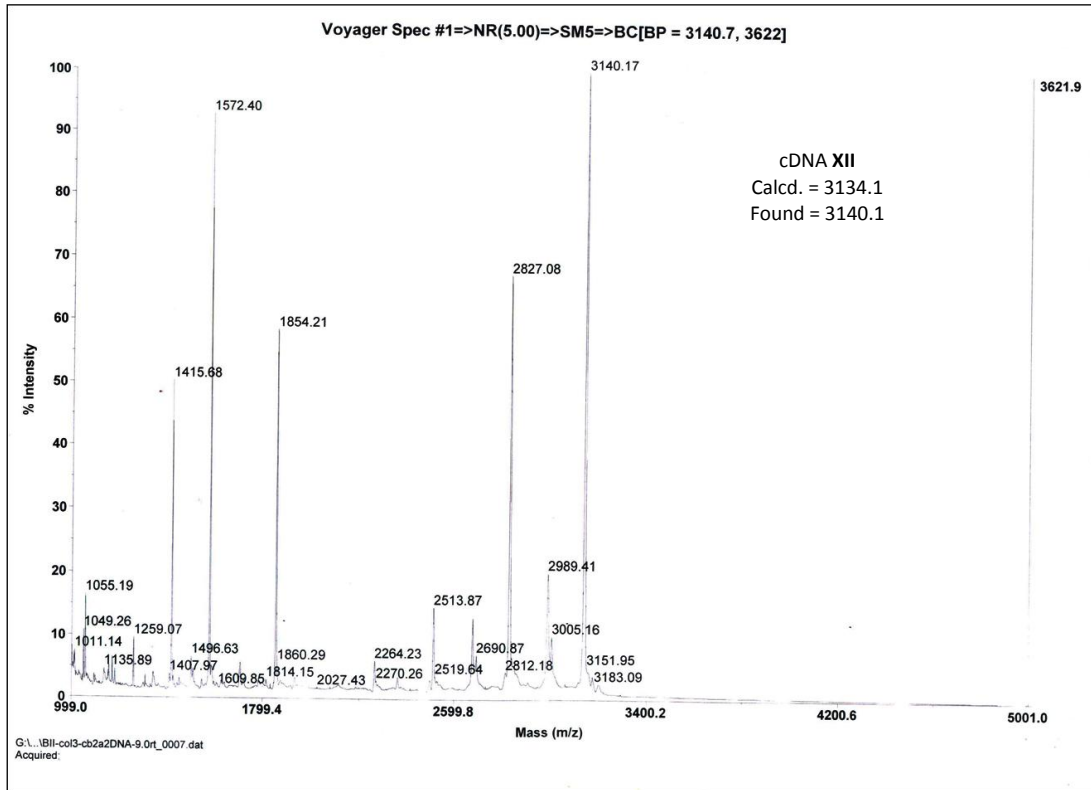
Chapter 3



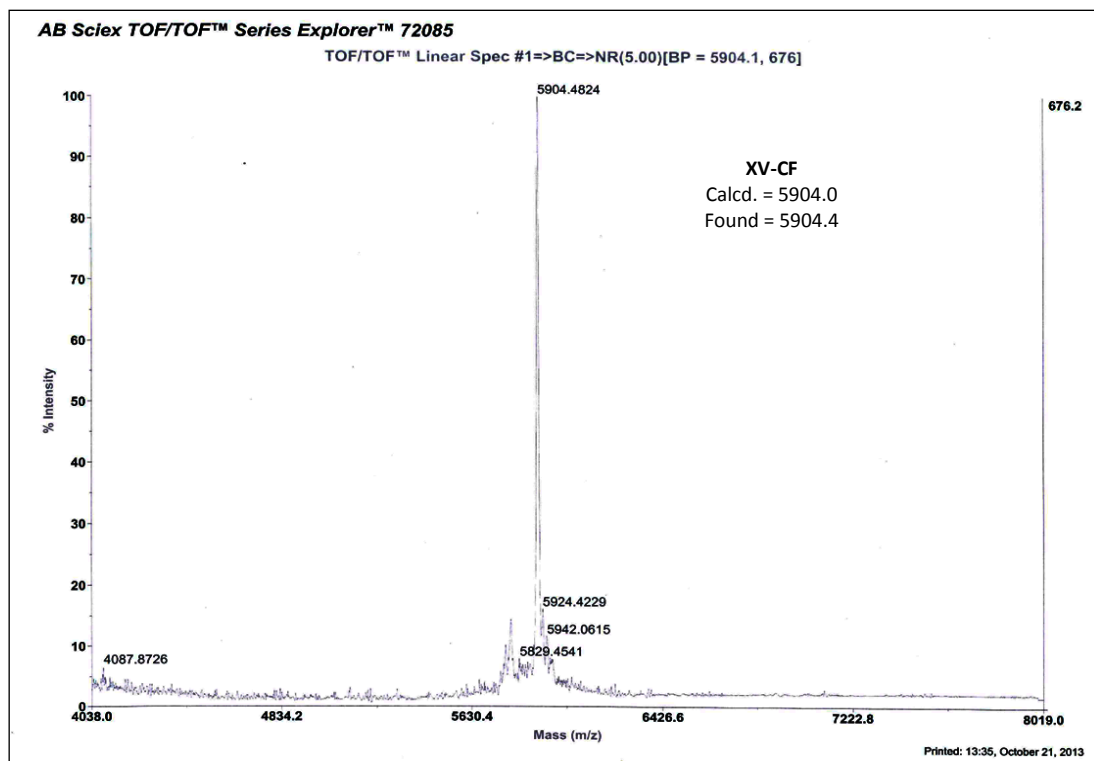
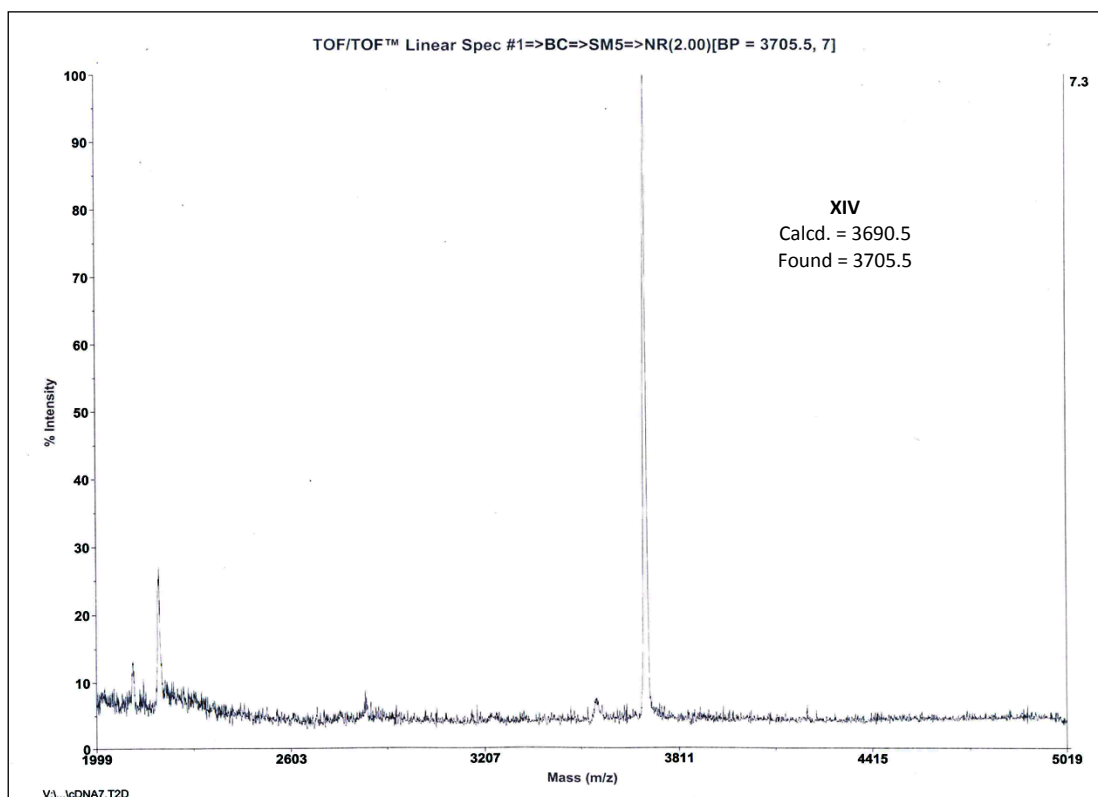
Chapter 3



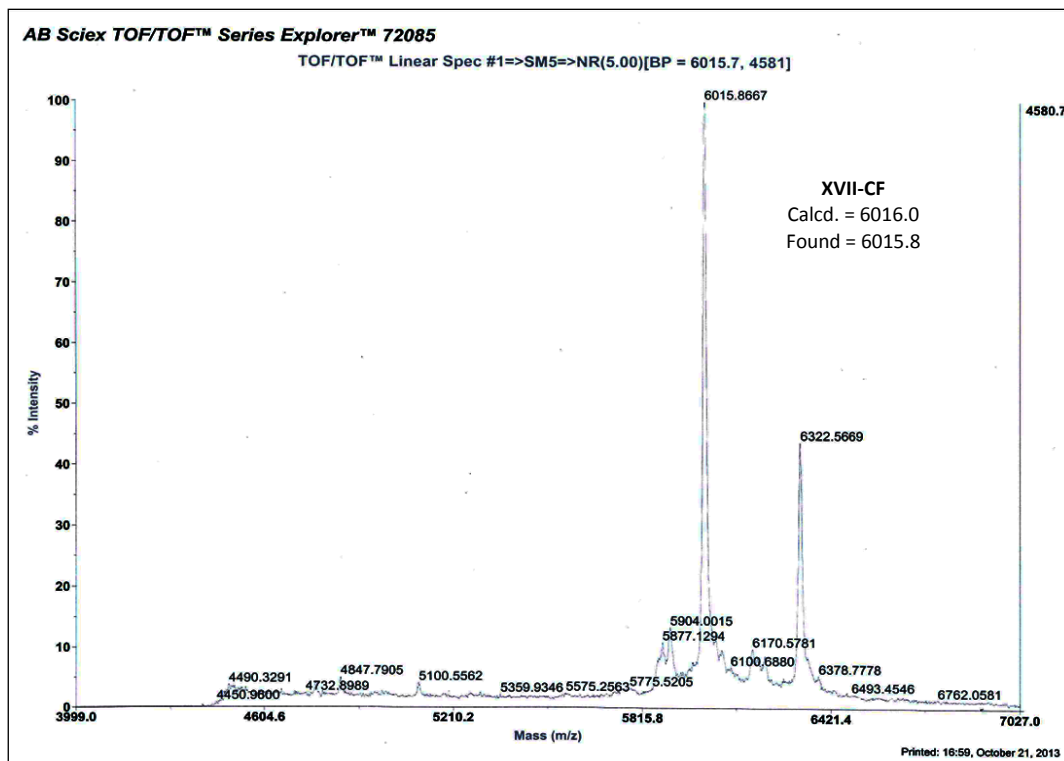
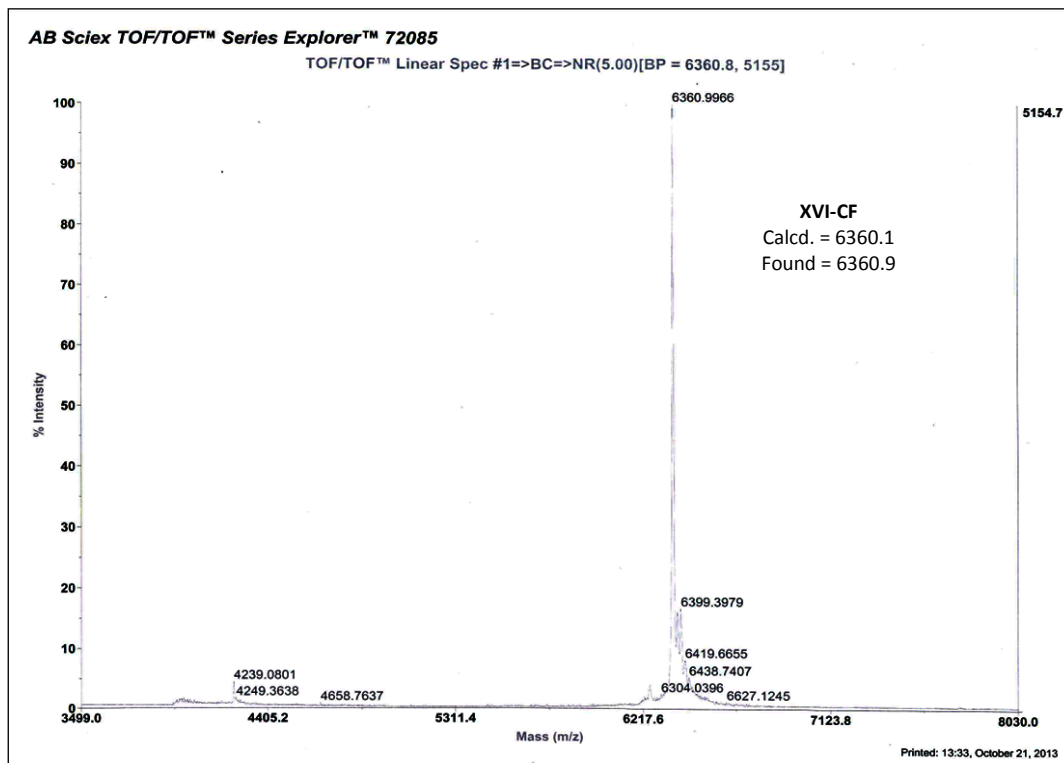
Chapter 3



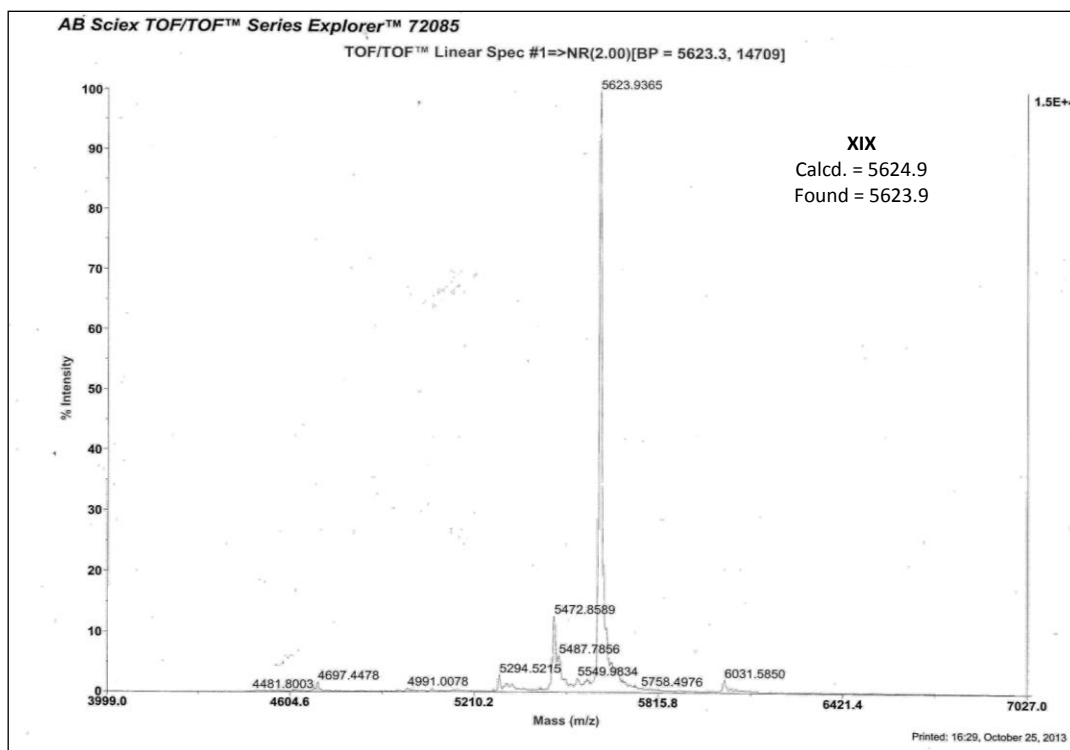
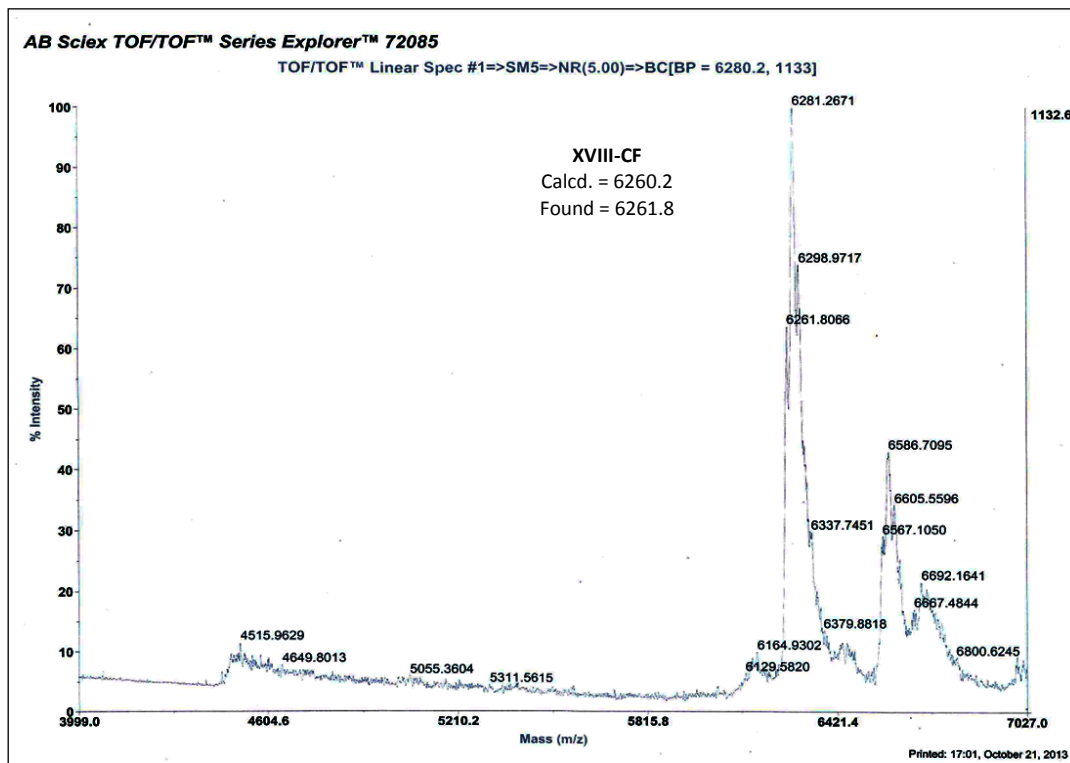
Chapter 3



Chapter 3



Chapter 3



3C References

1. (a)Nielsen, P. E.; Egholm, M.; Berg, R. H.; Buchardt, O., *Science* **1991**, 254, 1497-1500. (b)Hyrup, B.; Nielsen, P. E., *Bioorg. Med. Chem. Lett.* **1996**, 4, 5-23. (c)Goodchild, J., *Bioconjugate Chem.* **1990**, 1, 165-177.
2. (a)Metaferia, B.; Wei, J. S.; Song, Y. K.; Evangelista, J.; Aschenbach, K.; Johansson, P.; Wen, X.; Chen, Q.; Lee, A.; Hempel, H.; Gheeya, J. S.; Getty, S.; Gomez, R.; Khan, J., *PLoS ONE* **2013**, 8, e58870. (b)Esiobu, N., *Diagnostic Bacteriology Protocols:Methods in Molecular Biology* **2006**, 345, 131-140. (c)Shen, Y.; Shrestha, R.; Ibricevic, A.; Gunsten, S. P.; Welch, M. J.; Wooley, K. L.; Brody, S. L.; Taylor, J.-S. A.; Liu, Y., *Interface Focus* **2013**, 3, DOI:10.1098/rsfs.2012.0059.
3. Nielsen, P. E.; Haaima, G., *Chem. Soc. Rev.* **1997**, 26, 73-78.
4. Demidov, V. V.; Potaman, V. N.; Frank-Kamenetskii, M. D.; Egholm, M.; Buchardt, O.; Sonnichsen, S. H.; Nielsen, P. E., *Biochem. Pharmacol.* **1994**, 48, 1310-1313.
5. (a)Uhlmann, E.; Peyman, A.; Breipohl, G.; Will, D. W., *Angew. Chem. Int. Ed.* **1998**, 37, 2796-2823. (b)Uhlmann, E.; Will, D. W.; Breipohl, G.; Langer, D., *Angew. Chem. Int. Ed. Engl.* **1996**, 35, 2632-2635.
6. Rasmussen, H.; Kastrop, J. S.; Nielsen, J. N.; Nielsen, J. M.; Nielsen, P. E., *Nature Struct. Biol.* **1997**, 4, 98-101.
7. (a)Ganesh, K. N.; Nielsen, P. E., *Curr. Org. Chem.* **2000**, 4, 931-943. (b)Kumar, V. A., *Eur. J. Org. Chem.* **2002**, 2002, 2021-2032.
8. (a)Gait, M. J.; Jones, A. S.; Walker, R. T., *J. Chem. Soc., Perkin Trans.* **1974**, 1, 1684-1686. (b)Mungall, W. S.; Kaiser, J. K., *J. Org. Chem.* **1977**, 42, 703-706. (c)Coull, J. M.; Carlson, D. V.; Weith, H. L., *Tetrahedron Lett.* **1987**, 28, 745-748. (d)Stirchak, E. P.; Summerton, J. E.; Weller, D. D., *Nucleic Acids Res.* **1989**, 17, 6129-6141. (e)Prhavc, M.; Lesnik, E. A.; Mohan, V.; Manoharan, M., *Tetrahedron Lett.* **2001**, 42, 8777-8780.
9. Stirchak, E. P.; Summerton, J. E.; Weller, D. D., *J. Org. Chem.* **1987**, 52, 4202-4206.
10. Habus, I.; Tamsamani, J.; Agrawal, S., *Bioorg. Med. Chem. Lett.* **1994**, 8, 1065-1070.

Chapter 3

11. Habus, I.; Agrawal, S., *Nucleosides, Nucleotides and Nucleic Acids* **1995**, 14, 1853-1859.
12. Meena; Kumar, V. A., *Bioorg. Med. Chem. Lett.* **2003**, 11, 3393-3399.
13. Madhuri, V.; Kumar, V. A., *Org. Biomol. Chem.* **2010**, 8, 3734-3741.
14. Dueholm, K. L.; Egholm, M.; Behrens, C.; Christensen, L.; Hansen, H. F.; Vulpius, T.; Petersen, K. H.; Berg, R. H.; Nielsen, P. E.; Buchardt, O., *J. Org. Chem.* **1994**, 59, 5767-5773.
15. Haaima, G.; Lohse, A.; Buchardt, O.; Nielsen, P. E., *Angew. Chem. Int. Ed. Engl.* **1996**, 35, 1939-1942.
16. Merrifield, R. B., *J. Am. Chem. Soc.* **1963**, 85, 2149-2154.
17. Kaiser, E.; Colescott, R. L.; Bossinger, C. D.; Cook, P. I., *Anal. Biochem.* **1970**, 34, 595-598.
18. (a)Gait, M. J., *Oligonucleotide synthesis: A practical approach.* **1984**, IRL Press Oxford, UK 217. (b)Agrawal, S., *Protocols for oligonucleotides and analogs Synthesis and Properties Methods in Molecular Biology.* (ed): . **1993**, vol.20 Totowa, NJ,. Humana Press, Inc.
19. (a)Job, P., *Ann. Chim.* **1928**, 9, 113-203. (b)Cantor, C. R.; Schimmel, P. R., *Biophys. Chem. Part III* **1980**, 624.
20. Egholm, M.; Buchardt, O.; Christensen, L.; Behrens, C.; Frier, S.; Driver, D. A.; Berg, R. H.; Kim, S. K.; Norden, B.; Nielsen, P. E., *Nature* **1993**, 365, 566-568.
21. Gryaznov, S.; Skorski, T.; Cuco, C.; Nieborowska-Skorska, M.; Chiu, C. Y.; Liyod, D.; Chen, J.; Koziolkiewicz, M.; Calabretta, B., *Nucleic Acids Res.* **1996**, 24, 1508-1514.
22. Sambrook, J.; Fritsch, E. F.; Maniatis, T., In *Molecular Cloning: A Laboratory Manual*, 2. **1989**, Cold Spring Harbor Laboratory Press: Cold Spring Harbor, NY.
23. Venter, J. C.; Adams, M. D.; Sutton, G. G.; Kerlavage, A. R.; Smith, H. O.; Hunkapiller, M., *Science* **1998**, 280, 1540-1542.
24. Khakshoor, O.; Kool, E. T., *Chem. Commun.* **2011**, 47, 7018-7024.

Chapter 3

25. (a)Novoselov, K. S.; Geim, A. K.; Morozov, S. V.; Jiang, D.; Zhang, Y.; Dubonos, S. V.; Grigorieva, I. V.; Firsov, A. A., *Science* **2004**, 306, 666-669. (b)Geim, A. K.; Novoselov, K. S., *Nat. Mater.* **2007**, 6, 183-191.
26. Hummers, W. S.; Offeman, R. E., *J. Am. Chem. Soc.* **1958**, 80, 1339-1339.
27. Fan, X.; Peng, W.; Li, Y.; Li, X.; Wang, S.; Zhang, G.; Zhang, F., *Adv. Mater.* **2008**, 20, 4490-4493.
28. (a)Tang, Z.; Wu, H.; Cort, J.; Buchko, G.; Zhang, Y.; Shao, Y.; Aksay, I.; Liu, J.; Lin, Y., *Small* **2010**, 6, 1205-1209. (b)Li, F.; Y.Huang; Yang, Q.; Zhong, Z.; Li, D.; Wang, L.; Songa, S.; Fan, C., *Nanoscale* **2010**, 2, 1021-1026. (c)Wu, C.; Zhou, Y.; Miaoa, X.; Ling, L., *Analyst* **2011**, 136, 2106-2110. (d)Wu, M.; Kempaiah, R.; Huang, P.-J.; Maheshwari, V.; Liu, J., *Langmuir* **2011**, 27, 2731-2738. (e)Chung, C.; Kim, Y.-K.; Shin, D.; Ryoo, S.-R.; Hong, B.; Min, D.-H., *Acc.Chem. Res.* **2013**, DOI: 10.1021/ar300159f.
29. (a)Liu, Z.; Li, X.; Tabakman, S. M.; Jiang, K.; Fan, S.; Dai, H., *J. Am. Chem. Soc.* **2008**, 130, 13540-13541. (b)Chen, Y.; Liu, H.; Ye, T.; Kim, J.; Mao, C., *J. Am. Chem. Soc.* **2007**, 129, 8696-8697. (c)Heller, D. A.; Jin, H.; Martinez, B. M.; Patel, D.; Miller, B. M.; Yeung, T.-K.; Jena, P. V.; Hobartner, C.; Ha, T.; Silverman, S. K.; Strano, M. S., *Nat. Nanotechnol.* **2009**, 4, 114-120. (d)Chen, Z.; Tabakman, S. M.; Goodwin, A. P.; Kattah, M. G.; Darancioglu, D.; Wang, X.; Zhang, G.; Li, X.; Liu, Z.; Utz, P. J.; Jiang, K.; Fan, S.; Dai, H., *Nat. Biotechnol.* **2008**, 26, 1285 - 1292. (e)Sathishkumar, B. C.; Brown, L. O.; Gao, Y.; Wang, C.-C.; Wang, H.-L.; Doorn, S. K., *Nat. Nanotechnol.* **2007**, 2, 560-564.
30. (a)Cao, L.; Wang, X.; Meziani, M. J.; Lu, F.; Wang, H.; Luo, P. G.; Lin, Y.; Harruff, B. A.; Veca, L. M.; Murray, D.; Xie, S.-Y.; Sun, Y.-P., *J. Am. Chem. Soc.* **2007**, 129, 11318-11319. (b)Fu, C.-C.; Lee, H.-Y.; Chen, K.; Lim, T.-S.; Wu, H.-Y.; Lin, P.-K.; Wei, P.-K.; Tsao, P.-H.; Chang, H.-C.; Fann, W., *Proc. Natl. Acad. Sci. USA* **2007**, 104, 727-732. (c)Chang, I. P.; Hwang, K. C.; Chiang, C.-S., *J. Am. Chem. Soc.* **2008**, 130, 15476-15481.
31. Hao, C.; Ding, L.; Zhang, X.; Ju, H., *Anal. Chem.* **2007**, 79, 4442-4447.

Chapter 3

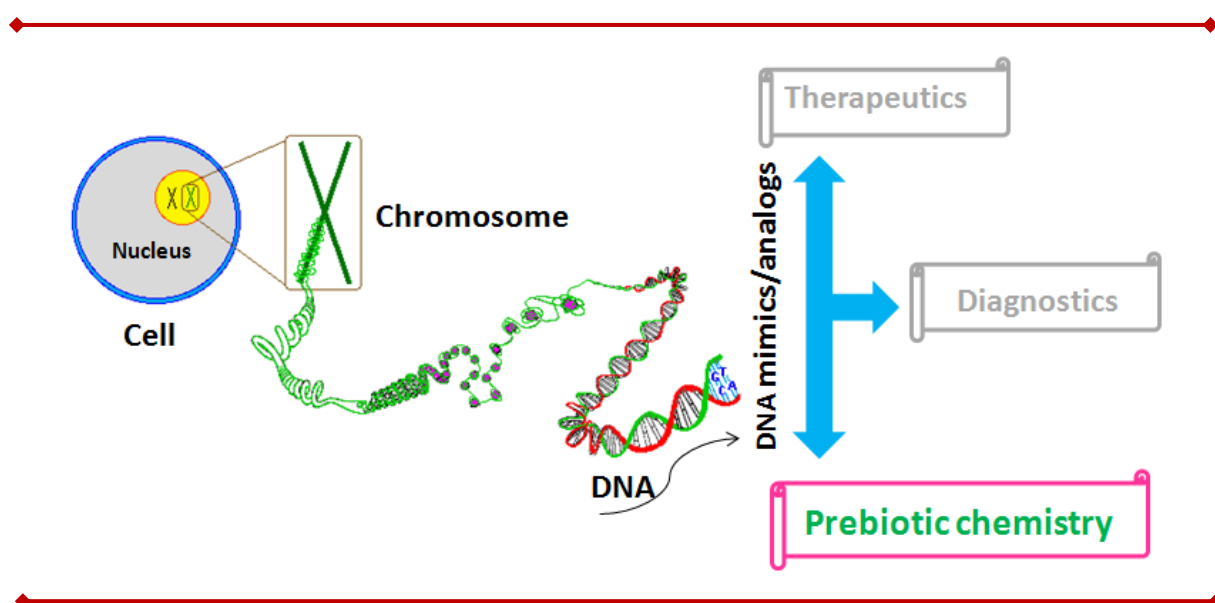
32. (a) Yang, R.; Jin, J.; Chen, Y.; Shao, N.; Kang, H.; Xiao, Z.; Tang, Z.; Wu, Y.; Zhu, Z.; Tan, W., *J. Am. Chem. Soc.* **2008**, 130, 8351-8358. (b) Yang, R.; Tang, Z.; Yan, J.; Kang, H.; Kim, Y.; Zhu, Z.; Tan, W., *Anal. Chem.* **2008**, 80, 7408-7413. (c) Zhu, Z.; Tang, Z.; Phillips, J. A.; Yang, R.; Wang, H.; Tan, W., *J. Am. Chem. Soc.* **2008**, 130, 10856-10857.
33. Mohanty, N.; Berry, V., *Nano Lett.* **2008**, 8, 4469-4476.
34. (a) Varghese, N.; Mogera, U.; Govindaraj, A.; Das, A.; Maiti, P. K.; Sood, A. K.; Rao, C. N. R., *Chem. Phys. Chem.* **2009**, 10, 206-210. (b) Liu, Z.; Robinson, J. T.; Sun, X.; Dai, H., *J. Am. Chem. Soc.* **2008**, 130, 10876-10877.
35. Kinali, M.; Arechavala-Gomez, V.; Feng, L.; Cirak, S.; Hunt, D.; Adkin, C.; Guglieri, M.; Ashton, E.; Abbs, S.; Nihoyannopoulos, P.; Garralda, M. E.; Rutherford, M.; McCulley, C.; Popplewell, L.; Graham, I. R.; Dickson, G.; Wood, M. J.; Wells, D. J.; Wilton, S. D.; Kole, R.; Straub, V.; Bushby, K.; Sewry, C.; Morgan, J. E.; Muntoni, F., Local restoration of dystrophin expression with the morpholino oligomer AVI-4658 in Duchenne muscular dystrophy: a single-blind, placebo-controlled, dose-escalation, proof-of-concept study. *Lancet Neurol.* **2009**, 8, 918-928.
36. He, S.; Song, B.; Li, D.; Zhu, C.; Qi, W.; Wen, Y.; Wang, L.; Song, S.; Fang, H.; Fan, C., *Adv. Funct. Mater.* **2010**, 20, 453-459.
37. Wu, M.; Kempaiah, R.; Huang, P.-J. J.; Maheshwari, V.; Liu, J., *Langmuir* **2011**, 27, 2731-2738.
38. Guo, Y.; Deng, L.; Li, J.; Guo, S.; Wang, E.; Dong, S., *ACS Nano* **2011**, 5, 1282-1290.
39. Huang, P.-J. J.; Liu, J., *Anal. Chem.* **2012**, 84, 4192-4198.
40. Kong, W. H.; Sung, D. K.; Kim, K. S.; Jung, H. S.; Gho, E. J.; Yun, S. H.; Hahn, S. K., *Biomaterials* **2012**, 33, 7556-7564.
41. Park, J. S.; Na, H.-K.; Min, D.-H.; Kim, D.-E., *Analyst* **2013**, 138, 1745-1749.
42. Ryoo, S.-R.; Lee, J.; Yeo, J.; Na, H.-K.; Kim, Y.-K.; Jang, H.; Lee, J. H.; Han, S. W.; Lee, Y.; Kim, V. N.; Min, D.-H., *ACS Nano* **2013**, 7, 5882-5891.
43. Wang, Y.; Li, Z.; Weber, T. J.; Hu, D.; Lin, C.-T.; Li, J.; Lin, Y., *Anal. Chem.* **2013**, 85, 6775-6782.

Chapter 3

44. Huang, P.-J. J.; Liu, J., *Nanomaterials* **2013**, 3, 221-228.
45. Guo, S.; Du, D.; Tang, L.; Ning, Y.; Yao, Q.; Zhang, G.-J., *Analyst* **2013**, 138, 3216-3220.
46. (a)Micklefield, J., *Curr. Med. Chem.* **2001**, 8, 1157-1179. (b)Kumar, V. A., *Eur. J. Org. Chem.* **2002**, 2002, 2021-2023. (c)Kurreck, J., *Eur. J. Biochem.* **2003**, 270, 1628-1644. (d)Kumar, V. A.; Ganesh, K. N., *Curr. Top. Med. Chem.* **2007**, 7, 715-726.
47. Micklitsch, C. M.; Oquare, B. Y.; Zhao, C.; Appella, D. H., *Anal. Chem.* **2013**, 85, 251-257.
48. Obernosterer, G.; Martinez, J.; Alenius, M., *Nature Protocols* **2007**, 2, 1508 -1514.
49. Várallyay, É.; Burgyán, J.; Havelda, Z., *Nature Protocols* **2008**, 3, 190-196.
50. Nazarenko, I.; Pires, R.; Lowe, B.; Obaidy, M.; Rashtchian, A., *Nucleic Acids res.* **2002**, 30, 2089-2195.
51. Patil, A. J.; Vickery, J. L.; Scott, T. B.; Mann, S., *Advanced Materials* **2009**, 21, 3159-3164.
52. Xiong, Z.; Zhang, L. L.; Ma, J.; Zhao, X. S., *Chem. Commun.* **2010**, 46, 6099-6101.
53. Debgupta, J.; Kakade, B.; Pillai, V. K., *Phys. Chem. Chem. Phys.* **2011**, 13, 14668-14674.
54. Nielsen, P. E.; Egholm, M., Peptide Nucleic Acids Protocols and Applications, Eds. *Horizon Scientific press* **1999**.
55. (a)Koshkin, A. A.; Fensholdt, J.; Pfundheller, H. M.; Lomholt, C., *J. Org. Chem.* **2001**, 66, 8504-8512. (b)Kotikam, V.; Kumar, V. A., *Tetrahedron* **2013**, 69, 6404-6408.
56. Kim, S. N.; Kuang, Z.; Slocik, J. M.; Jones, S. E.; Cui, Y.; Farmer, B. L.; McAlpine, M. C.; Naik, R. R., *J. Am. Chem. Soc.* **2011**, 133, 14480-14483.
57. Lee, K. H.; Oh, J. E., *Bioorg. Med. Chem.* **2000**, 8, 833-839.
58. Liu, B.; Sun, Z.; Zhang, X.; Liu, J., *Anal. Chem.* **2013**, 85, 7987-7993.

CHAPTER 4

Design and synthesis of erythrulose- based thymine and adenine monomers, and the attempted synthesis of modified oligonucleotides



4.1 Introduction

The two essential features that are required for the chemical emergence of life are the ability to store/transfer information as well as to perform selective chemical catalysis. Since the discovery of the double helical DNA structure, Watson-Crick base pairing has been widely speculated to be the likely mode of both information storage and communication in the earliest genetic polymers. The proposal by Woese, Crick and Orgel in the late 1960s,¹ of nucleic acids as the first biopolymers of life, gained interest by the discovery of catalytic RNA molecules in the early 1980s.² In spite of the discovery of many other RNA enzymes with functions relevant to the RNA world,³ it is still difficult to imagine RNA as the first genetic polymer, because it is difficult for an inherently unstable molecule as complex as RNA to self-assemble from the potential prebiotic molecules and to remain long enough in that environment to carry out the initial processes of information transfer and catalysis.⁴ These difficulties led to some suggestions that RNA is not the first genetic material but rather an intermediate/product during the evolution.⁵ Thus, if RNA is not the first genetic material, it may have been preceded by some other genetic polymers which could be synthesized easily by the same systems chemistry that gave rise to RNA. This hypothetical period in the evolution process is referred to as Pre-RNA world.⁶

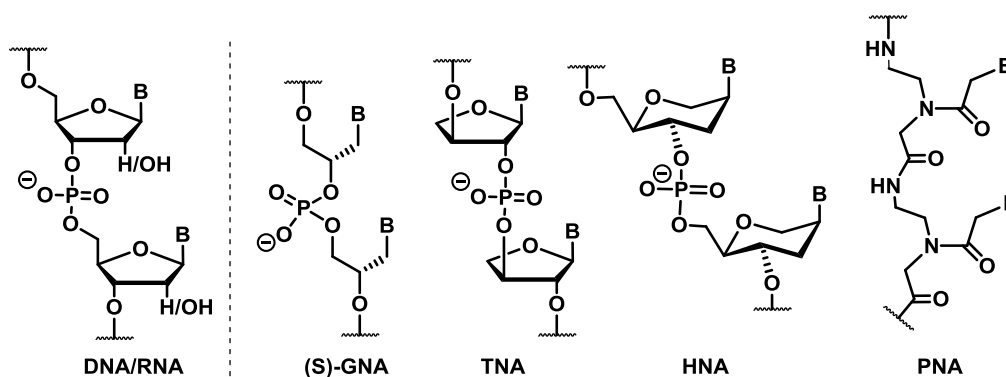


Figure 1 Chemical structures of natural genetic polymers (DNA/RNA) and recognized Pre-RNA synthetic genetic polymers.

Ground breaking experimental investigations by Eschenmoser and co-workers led to the discovery of α -L-threose nucleic acid (TNA),⁷ which emerged as the first synthetic genetic polymer with relevance in the Pre-RNA world. TNA is an artificial nucleic acid

polymer in which the natural five-carbon ribose sugar found in RNA was replaced with an unnatural four-carbon threose sugar, from the neighborhood of ribose sugar, and successive nucleosidic units were connected through 2' and 3' vicinal phosphodiester linkages (Figure 1). In spite of a repeating backbone unit shortened by one-atom, TNA still underwent informational Watson-Crick base pairing in an antiparallel strand orientation and also cross-paired with complementary strands of DNA and RNA.⁷

The ability of TNA to cross-pair with RNA along with the chemical simplicity of threose relative to ribose prompted careful consideration of TNA as a possible RNA progenitor.⁸ This consideration led the search for polymerases to examine the functional properties of TNA by *in vitro* selection. Primer-extension assays were performed to identify DNA polymerases that could copy DNA templates into TNA⁹ and copy TNA back into DNA.¹⁰ Once the enzymes were identified, further research was focused on developing the replication system for TNA that allowed pools of TNA molecules to be reverse-transcribed into DNA, amplified by PCR, and forward-transcribed back into TNA. These investigations led to the discovery of therminator DNA polymerase, an engineered variant of the 9° N DNA polymerase, as an efficient DNA-dependent TNA polymerase.¹¹ Though copying of long DNA to TNA was successful with the therminator polymerase, the attempts to copy a pool of DNAs to TNA failed.^{11a} This was recently achieved by Yu *et al.* by designing DNA libraries that either eliminated all G-residues in the template or presented G-nucleotides at sufficiently low frequency.¹² Using the hairpin strategy,¹³ it was also demonstrated that after several rounds of *in vitro* selection and amplification, TNA aptamers were evolved with high binding affinity and specificity to human thrombin. This demonstration also showed the TNA's ability to fold into tertiary structures with sophisticated chemical functions.¹² The similar selections are also possible to isolate the catalytic functioning novel TNA enzymes or “threozymes” by *in vitro* evolution.

Hexitol nucleic acid (HNA), developed by Herdewijn and co-workers as an antisense reagent¹⁴ was the second polymer to be studied in the context of synthetic genetics. HNA has a backbone structure composed of 1',5'-anhydrohexitol, having the nucleobase positioned at the 2' carbon and connected through the 4' and 6' phosphodiester

linkages (Figure 1). Unlike most other RNA analogs with six-membered carbohydrate moieties, such as homo-DNA and pyranosyl-RNA,¹⁵ HNA is capable of base pairing with complementary strands of itself and also RNA and DNA.¹⁶ HNA is also able to distinguish base pair mismatches more easily than natural oligonucleotides of the same sequence.^{16b} To explore the functional properties of HNA, Holliger and coworkers used a compartmentalized self-tagging strategy to evolve DNA polymerases capable of copying DNA templates into HNA.¹⁷ The authors constructed a library of polymerase variants based on the DNA enzyme *Thermococcus gorgonarius* (Tgo) polymerase and identified mutant polymerases capable of synthesizing HNA polymers on DNA templates. The screening of the polymerases produced enzymes that exhibited broad substrate specificity that could transcribe HNA, cyclohexenyl nucleic acids, locked nucleic acids (LNA), TNA, arabinose nucleic acids (ANA) and 2'-fluoro arabinose nucleic acids. Several RNA reverse transcriptases to reverse-transcribe the HNA polymers back into DNA were also constructed, from the vicinity of PolB family of DNA polymerases, and a new TgoT mutant, called RT521 was identified as an efficient HNA RT.¹⁷ Together, the two enzymes allowed for the *in vitro* evolution of HNA aptamers against HIV trans-activating response RNA and the protein hen egg lysozyme.

Peptide nucleic acids (PNAs, Figure 1) which were discovered in the context of gene targeting and gene therapeutic drugs (Chapter 1) were also considered for their merits in the context of origin of life. It is conceivable that the formation of PNAs would have been possible under prebiotic earth conditions due to the high chemical stability of the simple aminoethylglycine backbone in PNA. Indeed, Miller and co-workers showed the formation of PNA building blocks such as N-(2-aminoethyl)glycine and nucleobase acetic acids in primordial soup experiments.¹⁸ Liu and colleagues recently described an *in vitro* translation, selection, and amplification system for PNA.¹³ Since PNA is not a substrate for any known polymerase, conditions were developed for the nonenzymatic sequence specific oligomerization of tetramer and pentamer PNA building blocks on DNA templates.¹³ The authors performed six iterative cycles of translation, selection, and amplification for PNA molecules that could incorporate a biotinylated codon into their sequence. The selection

showed an overall enrichment of $>10^6$ -fold of PNA-encoding DNA templates on streptavidin-coated beads. This example suggests that it should be possible to isolate functional PNA aptamers by *in vitro* selection. Recently, Banack and co-workers showed that the Cyanobacteria which arose early in the earth's history produce the *N*-(2-aminoethyl)glycine backbone of the PNAs.¹⁹ These observations validated PNAs also as one of the potential progenitors of RNA.

Glycol nucleic acids (GNA) have an acyclic backbone of propylene glycol and were first described for the nucleoside derivative in 1971.²⁰ GNA has been synthesized as oligonucleotides where it showed considerable stability in an anti-parallel GNA–GNA duplex and follows the Watson–Crick base pairing rules.²¹ While GNA does not form stable duplexes with DNA, (*S*)-GNA (but not (*R*)-GNA) forms stable hetero-duplexes with RNA.²² Szostak and co-workers demonstrated that GNA oligonucleotides containing N2'-P3' phosphoramidate linkages (npGNA) can be assembled via a nonenzymatic, template-directed ligation of 3'-imidazole-activated-2'-amino GNA dinucleotides.²³ This suggested that sequence information transfer can be achieved nonenzymatically in a genetic system based on GNA oligonucleotides.

4.2 Design of erythrose-based nucleic acids and rationale

Several synthetic nucleic acids have been evaluated in the context of 'origin of life' as synthetic genetic polymers, but only few of them gained importance with their potential functions demonstrated in Darwinian evolution (enzymatically or non-enzymatically) such as sugar-based GNA (3-carbon), TNA (4-carbon) and HNA (6-carbon) and the non-sugar-based PNA. These synthetic genetic polymers gained interest not only because of their functions but also the simple chemistry that allowed their emergence to be easily imagined in the earth's prebiotic environment, unlike RNA.

Our interest in the field of antisense therapy and in the ongoing exploration of nucleic acids in the context of evolutionary chemistry, prompted us to design a new acyclic modification. Our design is based on the four carbon carbohydrate L-erythrose, a ketose, perhaps the first of its kind in the modified nucleic acids, in contrast to the aldose in

Eschemoser's TNA. In our design, the nucleobases are attached to the backbone through a methylene linker to the sp² carbon (carbonyl group of ketose) and individual units are linked through vicinal phosphodiester linkages between 3' and 4' of successive units (Figure 2). The backbone is shortened by one-atom and is comparable to that in TNA and GNA. In order to achieve the informational base-pairing with cDNA/cRNA, in our design, the nucleobase is attached through a methylene linker to the sp² carbon of the backbone, causing more rigidity compared to the GNA and more flexibility compared to the TNA, which is a cyclic modification. It is of interest to note that both TNA and *S*-GNA have been proven to form stable hetero-duplexes with RNA.

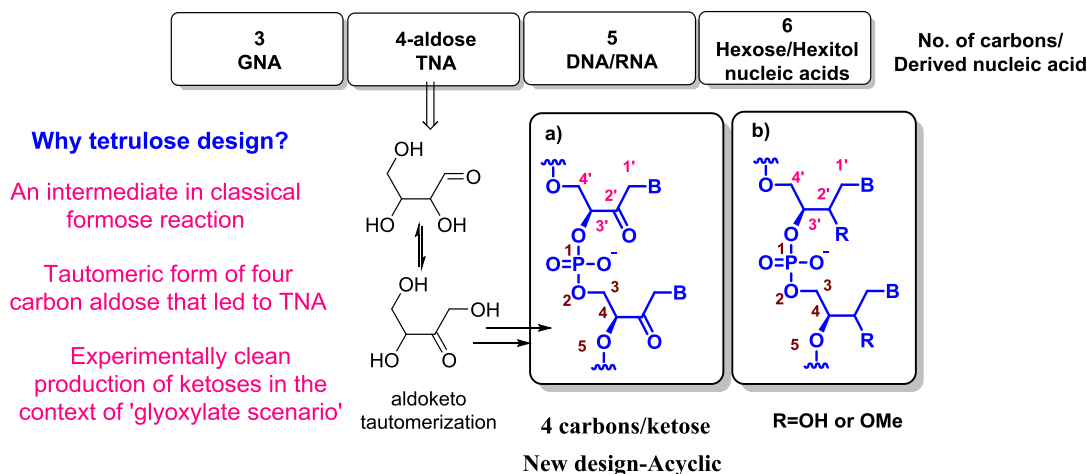


Figure 2 Design of 4-carbon ketose, erythrulose-based nucleic acids

The other factors that were considered in our design: (i) Erythrulose being also from the structural neighborhood of threose, it is likely that a mutant DNA polymerase could be found to copy the information from our designed nucleic acid to DNA, similar to the evolution of engineered DNA polymerases for copying information from TNA to DNA. (ii) Erythrulose is one of the branched sugars formed in the mixture during the classical formose reaction- the alkaline treatment of *para* formaldehyde. (iii) Erythrulose was chosen also based on the report of Sagi *et al.* where they reported the quantitative production of ketosugars with no detectable admixtures of aldoses from the reaction of glyoxylate with the small molecule aldehydes in water, in their attempt of exploration of the possible

prebiotic stock molecules in the ‘glyoxylate scenario’. (iv) Our designed modification derived from ketose might even help us in understanding nature’s selection of aldose-sugars rather than ketoses.

Our design also encompasses an acyclic analogue of RNA, where the carbonyl group was replaced by its reduced form, the hydroxyl group or its methyl derivative (Figure 2b). These mimics could also be achieved from a potential prebiotic molecule, tartaric acid, the abiotic formation of which, was proved via the cyanide-catalyzed dimerization of glyoxylate under alkaline conditions.

In this chapter, we present the synthesis of erythrose based-thymine and adenine monomers and also attempts towards synthesis of erythrose-containing oligonucleotides.

4.3 Retro synthetic analysis

The required phosphoramidite **1** could be achieved by the phosphorylation of free secondary hydroxy group subsequent to the selective dimethoxy tritylation of the primary hydroxyl in diol **2**. The diol **2** could be obtained from the deprotection of acetonide **3**. Prior to the nucleobase attachment, vicinal-diol of L-erythrose is protected as corresponding acetonide **4**.

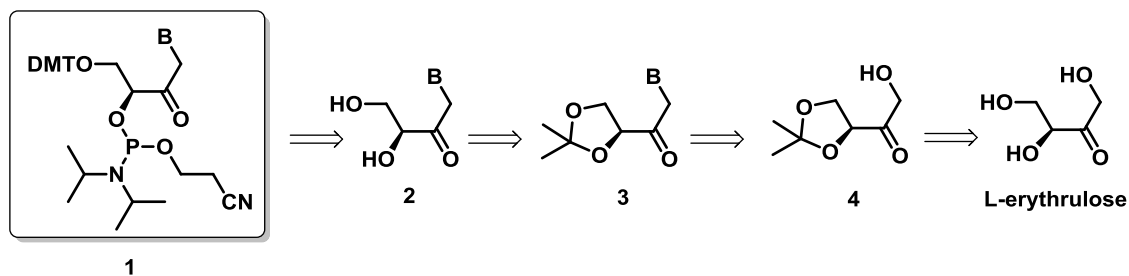
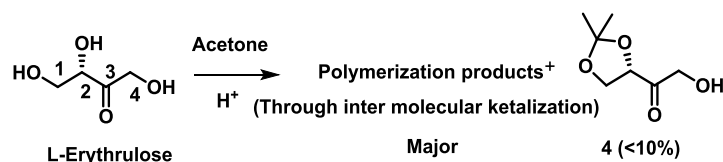


Figure 3 Retro synthetic analysis of the desired amidite from L-erythrose.

As is clear from the retrosynthetic analysis (Figure 3), the synthesis started with 1,2-acetonide protection of L-erythrose (Scheme 1). As in earlier literature reports, the yield of the acetonide protection was completely hampered by the formation of polymerization products. Several different dry reaction conditions were adopted but proved unsuccessful. As this first step of our synthetic route was low yielding, i.e. acetonide protection of L-

erythrose, we started looking for an alternative synthetic route to achieve our designed monomers.

Scheme 1 Acetonide protection of L-erythrose



4.4 Alternative retro synthetic analysis

Towards this end, we chose another natural chiral synthon, L-tartaric acid, a four carbon dicarboxylic acid having the required chirality at the hydroxyl groups in the backbone, as discussed in retrosynthetic analysis (Figure 4). The desired phosphoramidite **1** could be achieved by the phosphorylation of free secondary hydroxyl group subsequent to the selective dimethoxy tritylation of primary hydroxyl group of the diol **2**. The diol could be obtained by the pentyldiene deprotection **5**, where the carbonyl group could be derived from the oxidation of the secondary hydroxyl group that resulted from the debenzoylation of **6**. Prior to the nucleobase attachment, selective 1,2-hydroxyl groups of 1,2,4-triol (**8**) could be protected as pentyldiene **7**. The triol **8** could be derived from the AlCl_3 -mediated reduction of benzylidene dimethyltartrate **9**, which in turn could be obtained from the benzylidene protection of vicinal hydroxyl groups in the dimethyl-L-tartrate.

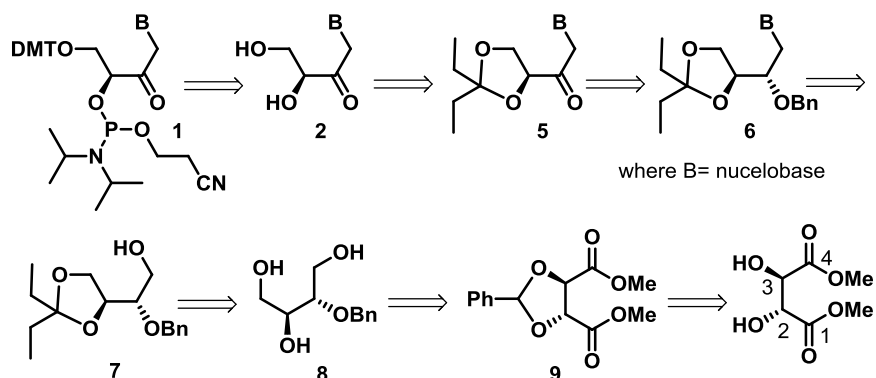
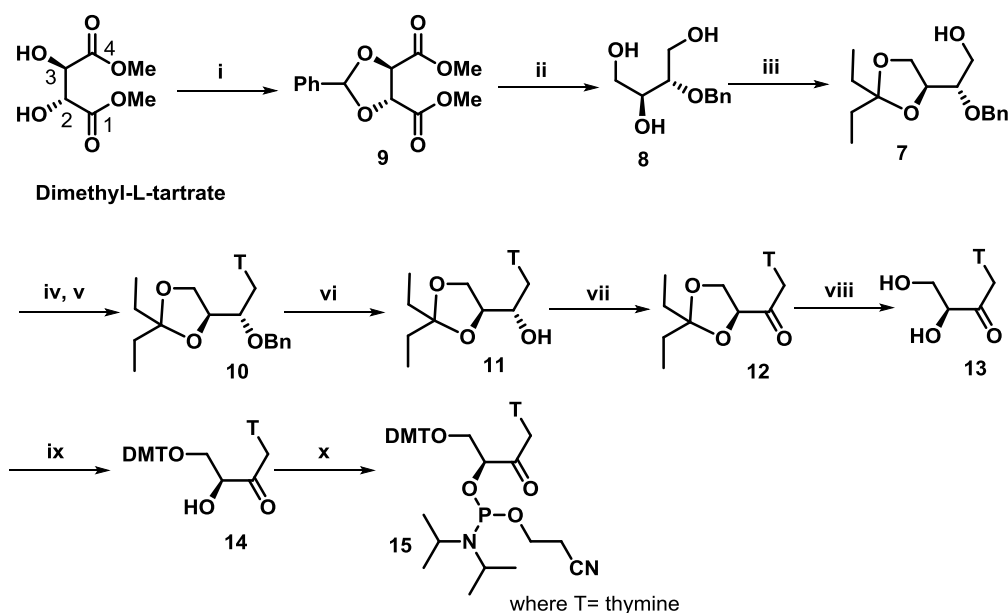


Figure 4 Retro synthetic analysis of the desired amidite from dimethyl-L-tartrate.

4.5 Synthesis of erythrulose- based monomers

The synthesis began with the benzylidene protection of 2,3-dihydroxy-dimethyltartrate using benzaldehydedimethylacetal under reduced pressure to furnish **9** (Scheme 1). Compound **9** was subjected to aluminium trichloride assisted LAH reduction to result 1,2,4-triol **8**. Pentylidene protection of the 1,2-diol of the 1,2,4-triol **8** was accomplished using 3-pentanone and *p*-toluenesulphonic acid in dry THF. The alcohol **7** was mesylated using mesyl chloride and triethylamine in dry DCM and further the mesylate was displaced with the nucleobase thymine to result in compound **10**. Debenzoylation of **10** was unsuccessful under the hydrogenation conditions using 10% Pd/C or Pd(OH)₂. Benzyl deprotection of compound **10** was achieved under ammonium formate mediated hydrogenation conditions to give the free hydroxy compound **11**.

Scheme 2 Synthesis of erythrulose-based thymine monomer from dimethyl-L-tartrate

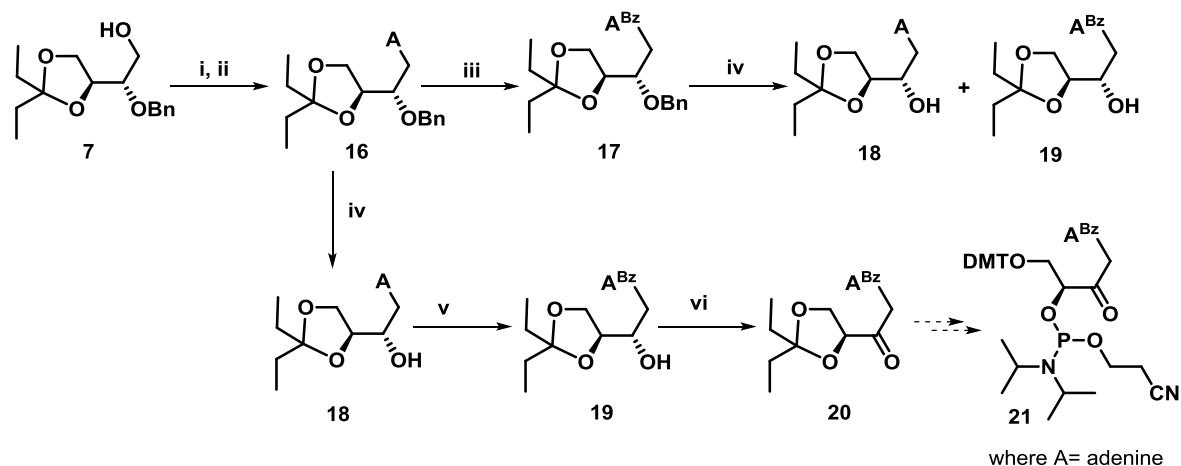


Reagents and conditions (i) PhCH(OMe)₂, under reduced pressure (ii) LiAlH₄, AlCl₃, Et₂O, dry DCM (iii) 3-pentanone, dry THF, *p*-TSA (iv) MsCl, Et₃N, dry DCM (v) NaH, dry DMF, thymine (vi) 10% Pd/C, ammonium formate, MeOH, reflux (vii) CrO₃, Ac₂O, Py, dry DCM (viii) 5% aq. H₂SO₄, MeOH (ix) DMT-Cl, Py, cat.DMAP (x) 2-cyanoethyl-N,N-diisopropylchlorophosphine, DIPEA, dry DCM.

The oxidation of the secondary hydroxyl group of **11** to its corresponding keto derivative was unsuccessful using the reagents IBX, DMP and $\text{SO}_3\cdot\text{Py}$ complex. Compound **11** was successfully oxidized to its keto derivative **12**, using chromium trioxide and acetic anhydride in pyridine. Deprotection of pentyldiene using aqueous H_2SO_4 in methanol gave **13**, followed by the dimethoxy tritylation of the primary alcohol group in **13** furnished compound **14**. The phosphitylation of the free secondary hydroxyl group in **14** delivered the required amidite monomer **15**, which could be incorporated using a DNA synthesizer into the modified oligonucleotides of our interest.

To get the adenine monomer **21**, the mesylate of alcohol **7** was displaced with the adenine to result compound **16**. *N,N*-Dibenzoyl protection of the exocyclic amine of adenine in compound **16** using benzoyl chloride, followed by the NH_4OH mediated hydrolysis for 30 min gave *N*-benzoyl-adenine derivative **17**. Debenzylation of **17** under ammonium formate mediated hydrogenation conditions gave **19** along with the compound **18**, resulting from the hydrolysis of *N*-benzoyladenine by the *in situ* generated ammonia in the reaction conditions.

Scheme 3 Synthesis of erythrose-based adenine monomer from dimethyl-L-tartrate



Reagents and conditions (i) MsCl , NEt_3 , dry DCM (ii) NaH , dry DMF, Adenine (iii) a) Bz-Cl , pyridine, rt, overnight b) NH_4OH , MeOH, rt, 30 min (iv) 10% Pd/C , ammonium formate, MeOH, reflux, 1 h (v) a) TMS-Cl , Py b) Bz-Cl , NH_4OH (vi) CrO_3 , Ac_2O , Py, dry DCM.

To circumvent this problem, we then changed the sequence of the reactions. The debenzoylation of **16** was performed under hydrogenation conditions using ammonium formate as hydrogen source, which resulted in alcohol **18**. The free hydroxyl group of alcohol **18** was protected as its transient trimethylsilyl ether, followed by the benzoyl protection of the exocyclic amine in adenine to give compound **19**. The alcohol **19** was oxidized to the corresponding keto compound **20** using chromium trioxide and acetic anhydride in pyridine. Similar to the thymine monomer synthesis, deprotection of pentylidene followed by the dimethoxytritylation of the primary hydroxyl group and finally the phosphitylation of the secondary hydroxyl group could deliver the adenine monomer **21** having benzoyl protection for the adenine exo-cyclic amine.

4.6 Attempted synthesis of oligonucleotides

In order to examine the stability offered by the new erythrulose-based acyclic modification, we incorporated thymine monomer **15** in the sequences (used in the preceding chapters) using an automated DNA synthesizer. Using universal solid support, we synthesized two different sequences of oligonucleotides containing single and double modified units of erythrulose-based thymine monomer (**T^{Ery}**, Table 1). Synthesized modified oligonucleotides **ON1-ON4** were purified RP-HPLC. Unfortunately the MALDI-TOF analysis of the modified oligonucleotides revealed mass difference of ~220-290 Daltons. During the attempted synthesis of fully modified homo-octa-thymidinyl sequence, we noticed that the coupling efficiency of thymine monomer **15** to itself was very low on automated DNA synthesizer, based on DMT- color intensity monitoring (observing the bright orange color during the detritylation step, which indicates the efficiency of the coupling).

The mass values could not be attributed by us to any possible modified oligonucleotides resulting from unwanted reactions (probably at the carbonyl of thymine monomer **15**) during the sequence of chemical reactions on automated DNA synthesizer. Never-the-less, the suitable protection of the carbonyl as in monomer **15/21** could lead to

the envisioned erythrulose-based oligonucleotides designed by us. This will be pursued as future directions of this work.

Table 1 Modified DNA sequences, their MALDI-TOF mass analyses

Sequences/Code	Mass	
	Calcd.	Obsd.
cctcttacctcagt $\underline{\mathbf{T}}^{\text{Ery}}\text{aca}$ ON1	5352.9	5116.2
cctcttacc $\underline{\mathbf{T}}^{\text{Ery}}\text{cagt}\underline{\mathbf{T}}^{\text{Ery}}\text{aca}$ ON2	5338.8	5116.4
cttcttcc $\underline{\mathbf{T}}^{\text{Ery}}\text{t}$ ON3	2904.4	2616.5
cttct $\underline{\mathbf{T}}^{\text{Ery}}\text{cc}\underline{\mathbf{T}}^{\text{Ery}}\text{t}$ ON4	2890.4	2615.9

4.7 Conclusion

- ✓ We have designed an erythrulose-based oligonucleotides, from a four-carbon ketose L-erythrulose, whose abiotic formations was experimentally proved in the context of ‘glyoxylate scenario’.
- ✓ Though the planned synthetic route from L-erythrulose gave a low yielding first reaction, we have successfully achieved the monomer synthesis from a four-carbon tartaric acid, whose abiotic formation was also proved recently.
- ✓ Unfortunately the MALDI-TOF mass of the modified oligonucleotides did not match with the calculated, probably suggesting the need for the carbonyl protections of our designed monomer before its incorporation on DNA synthesizer.
- ✓ Additionally, like 2'-OMe RNA analogue we can also use OMe-substitution instead of having carbonyl in our design, which is exactly an acyclic version of the 2'-OMe RNA having backbone shortened by one-atom.

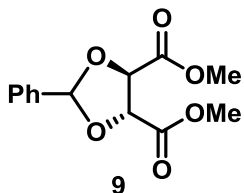
4.8 Experimental Section

General

All the non-aqueous reactions were carried out under the inert atmosphere of Nitrogen/Argon and the chemicals used were of laboratory or analytical grade. All solvents used were dried and distilled according to standard protocols. TLCs were carried out on pre-coated silica gel GF254 sheets (Merck 5554). All reactions were monitored by TLC and usual work-up implies sequential washing of the organic extract with water and brine followed by drying over anhydrous sodium sulfate and evaporation of the organic layer under vacuum. Column chromatographic separations were performed using silica gel 100-200 mesh (Merck) or 230-400 mesh (Merck) and using the solvent systems EtOAc/Petroleum ether or MeOH/DCM. ^1H and ^{13}C NMR spectra were obtained using Bruker AC-200, AC-400 or AC-500 NMR spectrometers. The chemical shifts are reported in delta (δ) values and referred to internal standard TMS for ^1H . LC-MS and HRMS-mass spectra were recorded on a Finnigan-Matt mass spectrometer. Modified DNA sequences were synthesized on Bioautomation Mer-Made 4 DNA synthesizer using standard β -cyanoethyl phosphoramidite chemistry. DNA sequences were purified using an increasing gradient of acetonitrile in 0.1N triethylammonium acetate of pH 7.0. The MALDI-TOF spectra were recorded on AB Sciex TOF/TOF series explorerTM 72085 instrument; THAP (2, 4, 6-trihydroxyacetophenone) was used as the matrix for modified oligonucleotides.

(2R,3R)-Dimethyl-2,3-O-benzylidene-L-tartrate, 9:

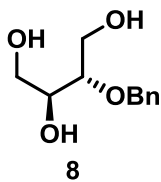
A mixture of dimethyl-(L)-tartrate (10.6 g, 58.2 mmol), benzaldehyde-dimethylacetal (44 mL, 291 mmol) and *p*-TSA (1.1 g, 5.8 mmol, 10 mol%) were stirred under reduced pressure for 8 h. The acid catalyst was neutralised with NaHCO₃ (3 g) and DCM was added to the reaction mixture and stirred for 1 h. The mixture was diluted with DCM, washed with water. The DCM layer was dried over anhydrous Na₂SO₄ and the solvent was removed under reduced pressure. Excess benzaldehyde-dimethylacetal and benzaldehyde were removed by reduced pressure distillation. The crude product was then recrystallised from EtOH and washed with EtOH to give **9** as white crystals in 87% yield.



¹H NMR (CDCl₃, 200 MHz): δ 3.17 (s, 3H), 3.81 (s, 3H), 4.84-4.98 (ABq, 2H, J = 3.92 Hz), 6.12 (s, 1H), 7.36-7.39 (m, 3H), 7.55-7.59 (m, 2H); ¹³C NMR (CDCl₃, 50 MHz): δ 52.3, 77.0, 76.6, 106.2, 126.8, 127.9, 129.5, 134.9, 169.0, 169.5.

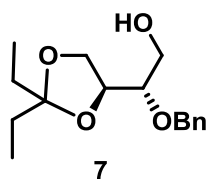
(2S,3S)-2-O-Benzyl-1,2,3,4-butanetetrol, 8:

Aluminium trichloride (13.11 g, 99.5 mmol) was dissolved in diethyl ether (65 mL) at -20 °C and to this was added LiAlH₄ (4.62 g, 100 mmol) in diethyl ether (75 mL) at -20 °C.



This mixture was then diluted with DCM (100 mL) and whilst maintained at -20 °C, compound **9** (4.5 g, 16.9 mmol) in DCM (100 mL) was added. The reaction mixture was allowed to warm to rt and then refluxed at 50 °C for 4 h. The reaction mixture was again cooled to -20 °C, then diluted with

THF (300 mL) and quenched with Na₂SO₄·10H₂O (45 g). This was then stirred at -20 °C for 2 h, allowed to warm to rt and left for 16 h. The products were filtered through a pad of celite. After removal of the solvents under reduced pressure the product was obtained by crystallisation from ethyl acetate and subjected to pentyldiene protection.

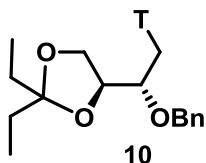
(2S,3S)-2-O-Benzyl-3,4-O-(3-pentyldiene)-1,2,3,4-butanetetrol, 7:

To a stirred solution of triol **8** (5 g, 23.5 mmol) in THF (60 ml) was added 3-pentanone (38 mL, 353.7 mmol) and *p*-toluenesulphonic acid (500 mg, 10% w/w). The resultant mixture was stirred at rt for 20 h.

NaHCO₃ solid was added to the reaction mixture until neutral pH. Filtration, removal of the solvents and purification by column chromatography (eluted in 15% EtOAc in petroleum ether) gave compound **7** as colourless oil in 76% yield along with the recovery of 9% triol **8**. ¹H NMR (CDCl₃, 200 MHz): δ 0.87 (m, 6H), 1.59-1.74 (m, 4H), 2.46 (m, 1H), 3.51-3.72 (m, 4H), 3.97-4.04 (m, 1H), 4.22-4.33 (m, 1H), 4.66 (ABq, 2H, J= 11.8 Hz), 7.29-7.38 (m, 5H); ¹³C NMR (CDCl₃, 50 MHz): δ 7.9, 8.1, 29.0, 29.5, 61.8, 66.1, 72.7, 77.2, 79.4, 113.2, 127.7, 127.8, 128.3, 138.2.

(2S,3S)-2-O-Benzyl-3,4-O-(3-pentylidene)-1-(N1-thyminy)-2,3,4-butanetriol, 10:

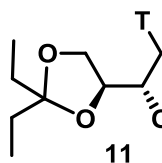
To a stirred solution of compound **7** (16.1 mmol, 4.5 g) in dry DCM (80 mL), Ms-Cl (32.3 mmol, 2.5 mL) and triethylamine (48.5 mmol, 6.7 mL) were added at 0



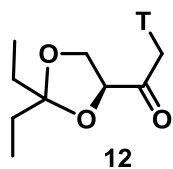
°C. The reaction mixture was stirred for 1 h at rt and quenched with 5% aq. KHSO₄ solution. The reaction mixture was diluted with DCM (300 mL) and was washed with water and brine solution, and the organic layer was dried over anhydrous Na₂SO₄. DCM was evaporated *in vacuo* which gave the crude mesylate as colorless oil. This crude product was directly subjected to the nucleobase attachment reaction. To a stirred mixture of thymine (32.3 mmol, 4.0 g) in dry DMF (40 mL) was added sodium hydride (60% suspension in mineral oil) (32.3 mmol, 1.3 g). To the reaction mixture, crude mesylate in DMF (10 mL) was added and stirred at 50 °C overnight. DMF was removed partially on rota evaporator under reduced pressure. The reaction mixture was diluted with EtOAc and water wash and brine solution wash were given. The organic layer was dried over anhydrous Na₂SO₄ and solvents were removed under vacuum. Column purification (eluted in 45%-50% EtOAc in petroleum ether) afforded compound **10** as a white solid in 64% yield over two steps, with the recovery of 13% un-reacted mesylate. ¹H NMR (CDCl₃, 200 MHz): δ 0.84 (dd, 6H, J= 15.4 Hz, 7.4 Hz), 1.57-1.74 (m, 4H), 1.8 (d, 3H, J= 0.88 Hz), 3.4 (m, 1H), 3.73-4.18 (m, 5H), 4.42-4.76 (ABq, 2H, J= 11.5 Hz), 6.98 (d, 1H, J= 1.14 Hz), 7.17-7.30 (m, 5H), 9.65 (bs, 1H); ¹³C NMR (CDCl₃, 50 MHz): δ 8.0, 11.9, 28.8, 29.4, 29.5, 49.4, 65.8, 73.9, 76.6, 76.8, 109.5, 113.4, 127.9, 128.2, 128.3, 137.2, 142.0, 150.9, 164.7; HRMS (EI): Mass calculated for C₂₁H₂₈O₅N₂Na (M+Na) 411.1890, found 411.1888.

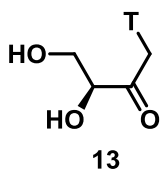
(2*S*,3*S*)-3,4-*O*-(3-pentylidene)-1-(*N*1-thyminy)-2,3,4-butanetriol, 11:

To a solution of compound **10** (7.21 mmol, 2.8 g) in MeOH (200 mL) were added 10% Pd/C (1.4 g, 50% w/W) and ammonium formate (72.1 mmol, 4.5 g). The reaction mixture was refluxed for 1h. The slurry was filtered through celite and the filtrate was evaporated under reduced pressure. The crude reaction mixture was column purified (eluted in 60% EtOAc in petroleum ether) to get compound **11** in 78% yield. ¹H NMR (CDCl₃, 200 MHz): δ 0.82 (dd, 6H, J= 14.6 Hz, 7.4 Hz), 1.57-1.72 (m, 4H), 1.86 (s, 3H), 3.33 (bs, 1H), 3.6-3.71 (m, 1H), 3.81-4.12 (m, 5H), 7.16 (d, 1H, J= 0.88 Hz), 9.98 (bs, 1H); ¹³C NMR (CDCl₃, 50 MHz): δ 7.9, 8.1, 12.1, 28.6, 29.2, 29.6, 51.4, 65.9, 69.6, 76.2, 109.8, 113.6, 142.2, 151.4, 164.8; HRMS (EI): Mass calculated for C₁₄H₂₂O₅N₂Na (M+Na) 321.1421, found 321.1420.

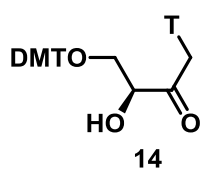
**(3*S*)-2-oxo-3,4-*O*-(3-pentylidene)-1-(*N*1-thyminy)-3,4-butanediol, 12:**

To CrO₃ (5 mmol, 0.5 g) in dry DCM (50 mL) were added acetic anhydride (5 mmol, 0.47 mL) and dry pyridine (10 mmol, 0.75 mL) and the resulting mixture stirred at rt for 5 min. Compound **11** (3.35 mmol, 1 g) was added to the reaction mixture and stirring continued at rt for 3h. As TLC analysis indicated the reaction to be incomplete, another equivalent of the reagents (CrO₃, Ac₂O, pyridine) were added and the resulting mixture stirred for a further 3 h. The reaction mixture was filtered through a celite bed and the DCM was removed under reduced pressure. The crude compound was column purified (eluted in 45%-50% EtOAc in petroleum ether) to furnish compound **12** as gummy liquid in 69% yield. ¹H NMR (CDCl₃, 400 MHz): δ 0.9 (t, 3H, J= 7.58 Hz), 0.97 (t, 3H, J= 7.34 Hz), 1.63 (q, 2H, J= 14.9, 7.34 Hz), 1.73 (ddd, 2H, J= 15.1, 7.5, 2.4 Hz), 1.90 (s, 3H), 4.04-4.08 (m, 1H), 4.22-4.26 (m, 1H), 4.56-4.6 (m, 1H), 4.66-4.86 (ABq, 2H, J= 18.5 Hz), 6.83 (d, 1H, J= 1.22 Hz), 9.55 (bs, 1H); ¹³C NMR (CDCl₃, 50 MHz): δ 8.0, 8.2, 12.2, 27.9, 28.9, 29.6, 53.8, 66.4, 79.4, 110.9, 115.6, 140.4, 150.8, 164.4, 202.8; HRMS (EI): Mass calculated for C₁₄H₂₁O₅N₂ (M+H) 297.1445, found 297.1447.



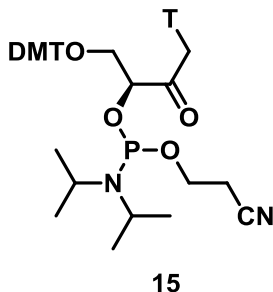
(3S)-2-oxo-1-(N1-thyminyl)-3,4-butanediol, 13:

Compound **12** (3.2 mmol, 0.95 g) was dissolved in MeOH (50 mL) followed by the addition of 2% aq. H₂SO₄ (10 mL). The resulting mixture was stirred at rt for 5h. When the reaction was completed, K₂CO₃ (5 g) solid was added, the solids were filtered off and the solvents were removed under reduced pressure. The crude compound was column purified (eluted in 95% EtOAc in petroleum ether) to get compound **13** in 87% yield. ¹H NMR (Acetone-d₆, 200 MHz): δ 1.81 (d, 3H, J= 1.03 Hz), 4.07-4.3 (m, 3H), 4.64-4.74 (m, 1H), 4.84 (d, 2H, J= 8.8 Hz), 4.9-4.97 (m, 1H), 7.34 (d, 1H, J= 1.15 Hz); ¹³C NMR (Acetone-d₆, 100 MHz): δ 12.1, 55.4, 64.2, 77.9, 110.3, 143.0, 152.1, 165.6, 206.9; HRMS (EI): Mass calculated for C₉H₁₂O₅N₂Na (M+Na) 251.0638, found 251.0642.

(3S)-4-O-DMT-2-oxo-1-(N1-thyminyl)-3,4-butanediol, 14:

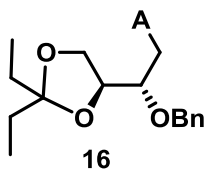
Compound **13** (2.6 mmol, 0.6 g) was dissolved in dry pyridine (15 mL). DMT-Cl (2.9 mmol, 0.98 g) and a catalytic amount of DMAP (~20 mg) were added. The reaction mixture was stirred at room temperature for 5–6 h. Pyridine was removed under reduced pressure and the residue dissolved in EtOAc. 10% aqueous NaHCO₃, water and brine washes were given to the organic layer. The organic layer was dried over anhydrous Na₂SO₄ and concentrated to dryness. The crude compound was column purified (eluted in 70% EtOAc in petroleum ether) to afford **14** in 72% yield as a white foam. ¹H NMR (CDCl₃, 200 MHz): δ 1.79 (s, 3H), 3.32 (dd, 1H, J= 10.1, 3.4 Hz), 3.61 (dd, 1H, J= 10.1, 3.0 Hz), 3.78 (s, 6H), 4.34-4.42 (m, 2H), 4.74 (m, 1H), 6.59 (s, 1H), 6.82 (m, 4H), 7.26-7.38 (m, 10H), 9.16 (bs, 1H); ¹³C NMR (Acetone-d₆, 50 MHz): δ 12.3, 55.5, 66.7, 76.8, 87.1, 110.1, 113.6, 113.9, 127.6, 128.3, 128.6, 128.7, 129.0, 130.0, 131.02, 131.06, 136.4, 136.7, 142.1, 145.9, 151.9, 159.6, 165.0, 206.1; HRMS (EI): Mass calculated for C₃₀H₃₀O₇N₂Na (M+Na) 553.1945, found 553.1954.

Phosphoramidite, 15:



To a solution of **14** (0.56 mmol, 0.3 g) in dry DCM (8 mL) was added DIPEA (2.5 mmol, 0.43 mL). 2-cyanoethyl-*N,N*-diisopropyl-chloro phosphine (1.2 mmol, 0.26 mL) was added to the solution at 0 °C and the reaction mixture was stirred at room temperature for 3 h. The contents were then diluted with DCM and washed with 5% NaHCO₃ solution. The organic phase was dried over anhydrous Na₂SO₄ and concentrated to foam. The residue was re-dissolved in DCM and precipitated with n-hexane to yield the corresponding phosphoramidite **15** in 76% as white foam. ³¹P NMR (Acetonitrile, D₂O as external standard, 160 MHz): δ 150.58, 152.21. HRMS (EI): Mass calculated for C₃₉H₄₇N₄O₈PNa (M+Na) 753.3024, found 753.3024.

(2*S*,3*S*)-2-*O*-Benzyl-3,4-*O*-(3-pentylidene)-1-(*N*9-adeninyl)-2,3,4-butanetriol, 16:

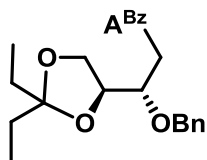


To a stirred solution of compound **7** (10.71 mmol, 3.0 g) in dry DCM (60 mL), Ms-Cl (21.4 mmol, 1.6 mL) and triethylamine (32.1 mmol, 4.4 mL) were added at 0 °C. The reaction mixture was stirred for 1 h at rt and quenched with 5% aq. KHSO₄ solution. The reaction mixture was diluted with DCM (250 mL) and was washed with water and brine solution, the organic layer was dried over anhydrous Na₂SO₄, and DCM was evaporated *in vacuo*, which gave the crude mesylate as colorless oil. This crude product was directly subjected to the nucleobase attachment reaction. To a stirring mixture of adenine (16.0 mmol, 2.1 g) in dry DMF (30 mL) was added sodium hydride (60% suspension in mineral oil) (16.0 mmol, 0.65 g). To the reaction mixture, crude mesylate in DMF (5 mL) was added and stirred at 50 °C overnight. DMF was removed partially on rota evaporator under reduced pressure. The reaction mixture was diluted with EtOAc and water wash and brine solution wash were given. The organic layer was dried over anhydrous Na₂SO₄ and solvents were removed under vacuum. Column purification (eluted in 90% EtOAc in petroleum ether) gave compound **16** as a white solid in 86% yield over two steps. ¹H NMR (CDCl₃, 200 MHz): δ 0.83-0.96 (m, 6H), 1.55-1.75 (m, 4H), 3.80-3.92 (m, 2H), 4.03-4.29 (m, 4H), 4.3 (ABq, 2H, J= 11.4 Hz), 6.57 (s, 2H), 7.04-7.2 (m, 5H), 7.82 (s, 1H), 8.29 (s, 1H); ¹³C NMR (CDCl₃,

50 MHz): δ 8.1, 28.7, 29.3, 44.5, 65.8, 73.5, 76.4, 113.5, 119.1, 127.7, 127.9, 128.1, 137.0, 141.4, 149.7, 152.7, 155.6; **HRMS (EI)**: Mass calculated for $C_{21}H_{28}O_3N_5$ (M+H) 398.2187, found 398.2191.

(2*S*,3*S*)-2-*O*-Benzyl-3,4-*O*-(3-pentylidene)-1-(*N*9-adeninyl^{Bz})-2,3,4-butanetriol, 17:

To the compound **16** (11 mmol, 4.3 g) in dry pyridine (40 mL) was added benzoyl chloride



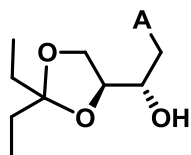
17

(33.0 mmol, 3.86 mL) and the resulting mixture was stirred at rt overnight. Pyridine was removed under reduced pressure and the reaction mixture was diluted with EtOAc. To the organic layer water and brine washes were given. EtOAc was removed on rota evaporator,

the crude compound dissolved in MeOH (30 mL) and 10% aq. NH_3 (15 mL) was added. the reaction mixture was stirred for 30 min. When the TLC analysis of the reaction mixture indicated the disappearance of dibenzoyl compound **16**, solvents were removed under reduced pressure. The crude reaction mixture was extracted with EtOAc and water and brine washes were given to the organic layer. The organic layer was dried over anhydrous Na_2SO_4 and solvents were removed under vacuum. Column purification (eluted in 60% EtOAc in petroleum ether) furnished compound **17** in 92% yield. **1H NMR** ($CDCl_3$, 200 MHz): δ 0.83-0.97 (m, 6H), 1.55-1.76 (m, 4H), 3.82-3.93 (m, 2H), 4.05-4.63 (m, 6H), 7.01-7.17 (m, 5H), 7.43-7.59 (m, 3H), 8.01 (s, 2H), 8.04 (s, 1H), 8.69 (s, 1H), 9.42 (bs, 1H); **^{13}C NMR** ($CDCl_3$, 50 MHz): δ 8.0, 28.5, 29.2, 44.5, 65.6, 73.3, 76.0, 76.1, 113.5, 122.5, 127.8, 128.0, 128.5, 132.4, 133.5, 136.7, 143.9, 149.2, 151.7, 152.1, 164.7.

(2*S*,3*S*)-3,4-*O*-(3-pentylidene)-1-(*N*9-adeninyl)-2,3,4-butanetriol, 18:

To a solution of compound **16** (7.0 mmol, 2.8 g) in MeOH (220 mL) was added 10% Pd/C



18

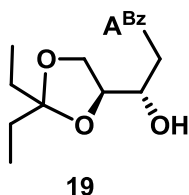
(65%-70% w/w, 1.8 g) and ammonium formate (70.5 mmol, 4.44 g). The reaction mixture was refluxed for 1h. The slurry was filtered through celite and the filtrate was evaporated under reduced pressure. Compound **18** was crystallized from MeOH as white crystals in 93% yield. **1H NMR**

(DMSO- d_6 , 200 MHz): δ 0.78-0.9 (m, 6H), 1.49-1.67 (m, 4H), 3.7-3.81 (m, 1H), 3.91-4.28 (m, 5H), 5.33 (s, 1H), 7.23 (s, 2H), 8.07 (s, 1H), 8.15 (s, 1H); **^{13}C NMR** (DMSO- d_6 , 50

MHz): δ 8.0, 8.1, 28.6, 29.0, 46.1, 65.3, 68.3, 77.0, 112.2, 118.6, 141.7, 149.6, 152.3, 155.9; **HRMS (EI)**: Mass calculated for $C_{14}H_{22}O_3N_5$ (M+H) 308.1717, found 308.1717.

(2*S*,3*S*)-3,4-*O*-(3-pentylidene)-1-(*N*9-adeninyl^{Bz})-2,3,4-butanetriol, 19:

To a solution of **18** in dry pyridine at 0 °C, was added dropwise TMS-Cl. After the mixture was stirred for 1 h, benzoyl chloride was added slowly. The reaction mixture was then stirred overnight. The reaction mixture was cooled and quenched with water. After stirring

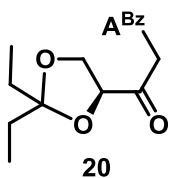


the mixture for 5 min, 25% NH_4OH solution was added. The resulting mixture was stirred at rt for 15 min and then concentrated. The residue was dissolved in methanol and 25% NH_4OH solution added and stirred further at rt for 30 min. After concentration, the crude residue was

purified by column chromatography (eluted in 80%-90% EtOAc in petroleum ether) to result in compound **19** in 90% yield. **¹H NMR** (Acetone- d_6 , 200 MHz): δ 0.79-0.92 (m, 6H), 1.51-1.7 (m, 4H), 3.87-4.17 (m, 4H), 4.32-4.54 (m, 2H), 7.48-7.66 (m, 3H), 8.1-8.14 (m, 2H), 8.29 (s, 1H), 8.59 (s, 1H); **¹³C NMR** (Acetone- d_6 , 100 MHz): δ 8.5, 8.53, 29.5, 30.1, 47.7, 66.5, 69.8, 77.9, 113.7, 124.8, 129.2, 129.4, 133.2, 145.9; **HRMS (EI)**: Mass calculated for $C_{21}H_{26}O_4N_5$ (M+H) 412.1979, found 412.1983.

(3*S*)-2-oxo-3,4-*O*-(3-pentylidene)-1-(*N*9-adeninyl^{Bz})-3,4-butanediol, 20:

To CrO_3 (9.48 mmol, 0.95 g) in dry DCM (50 mL) were added acetic anhydride (9.48 mmol, 0.9 mL) and dry pyridine (18.9 mmol, 1.45 mL) and the resulting mixture stirred at rt for 5 min. Compound **19** (3.35 mmol, 1 g) was added to



the reaction mixture and stirring continued at rt for 3h. As TLC analysis indicated the reaction to be incomplete, another equivalent of the reagents

(CrO_3 , Ac_2O , pyridine) was added and the resulting mixture stirred for a further 3 h. The reaction mixture was filtered through a celite bed and the DCM was removed under reduced pressure. The crude compound was column purified (eluted in 80%-90% EtOAc in petroleum ether) to furnish compound **20** as gummy liquid in 46% yield. **¹H NMR** (Acetone- d_6 , D_2O , 400 MHz): δ 0.88 (t, 3H, $J= 7.34$ Hz), 0.97 (t, 3H, $J= 7.58$ Hz), 1.67 (q, 2H, $J= 14.92$ Hz, 7.34 Hz), 1.78 (q, 2H, $J= 14.92$ Hz, 7.34 Hz), 4.02-4.16 (m, 3H), 4.3 (t,

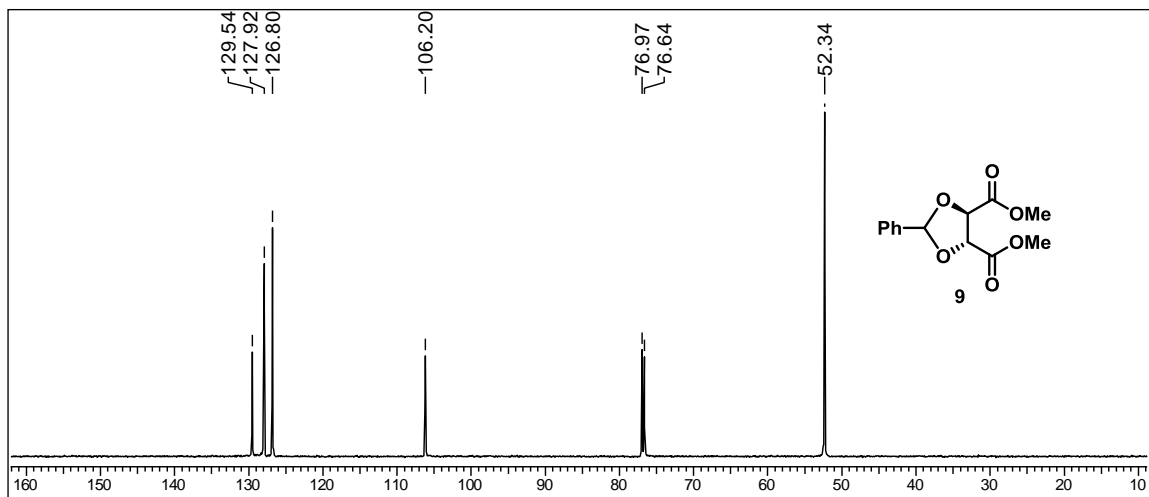
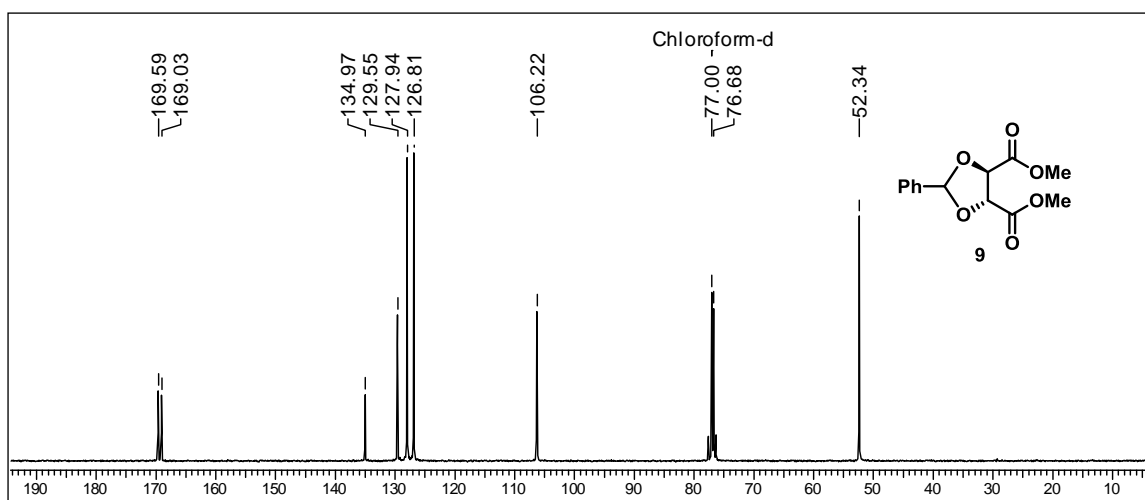
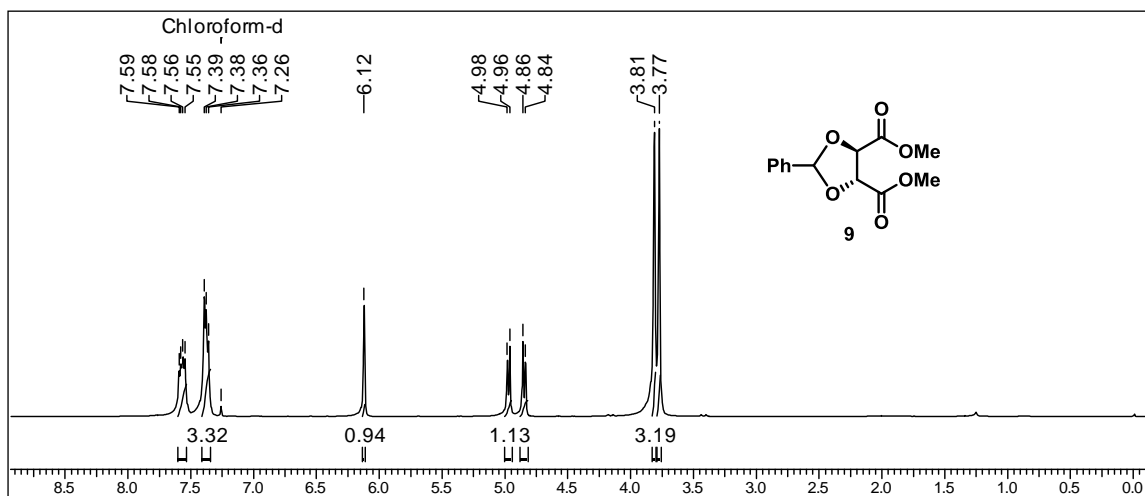
Chapter 4

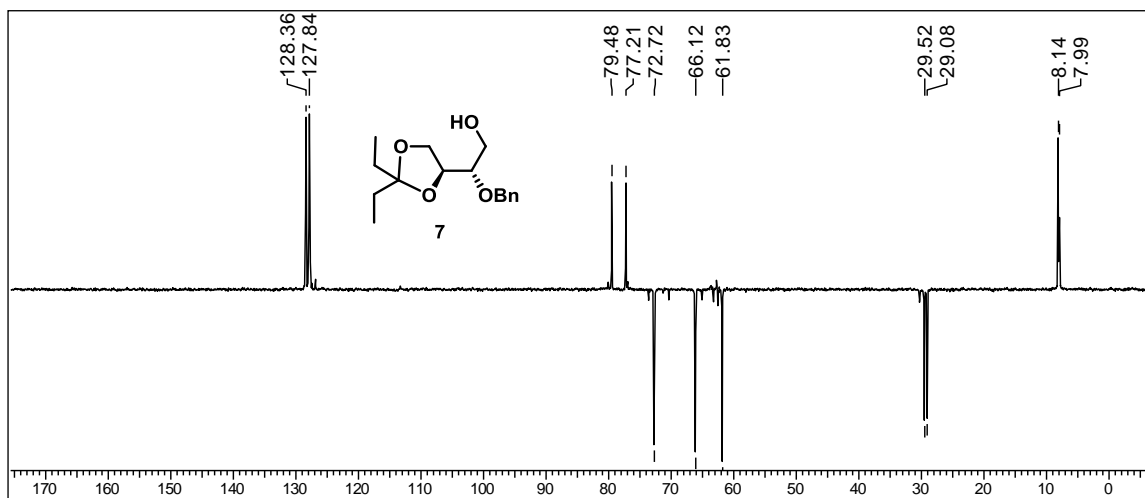
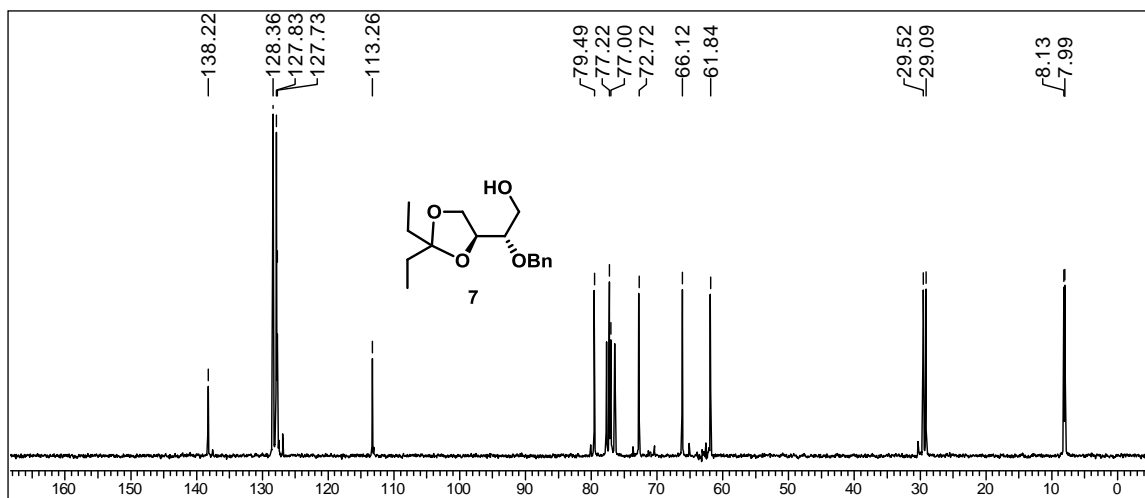
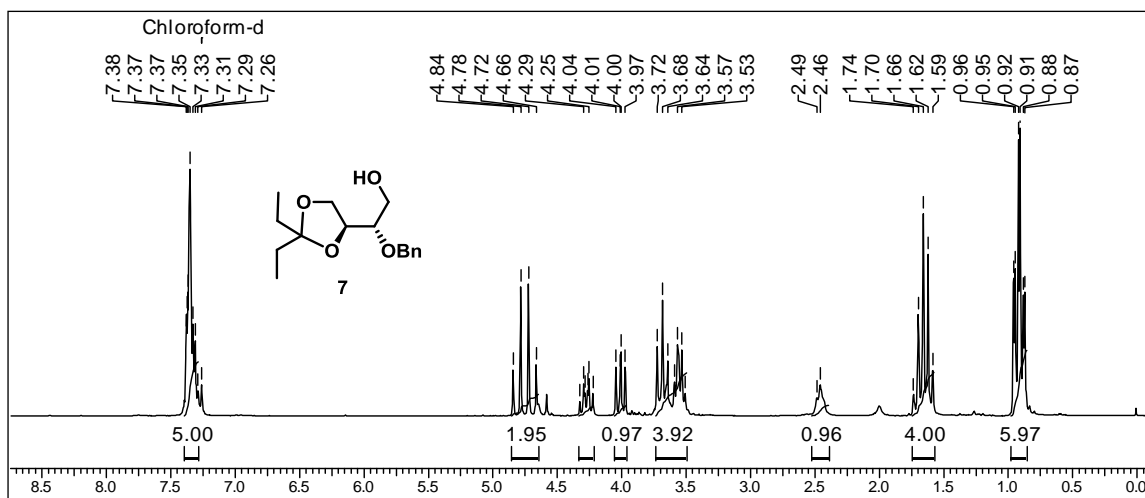
1H, J= 8.31 Hz), 4.85 (t, 1H, J= 6.36Hz), 7.49-7.6 (m, 3H), 8.1 (d, 2H, J= 7.34 Hz), 8.37 (s, 1H), 8.64 (s, 1H), 8.67 (bs, 1H); ¹³C NMR (Acetone-d₆, D₂O, 100 MHz): δ 8.3, 8.5, 28.5, 29.5, 50.5, 66.8, 80.0, 115.9, 123.7, 129.1, 129.3, 133.3, 134.6, 145.8, 150.5, 152.6, 153.1, 166.5, 203.1; **HRMS (EI)**: Mass calculated for C₂₁H₂₄O₄N₅ (M+H) 410.1823, found 410.1828.

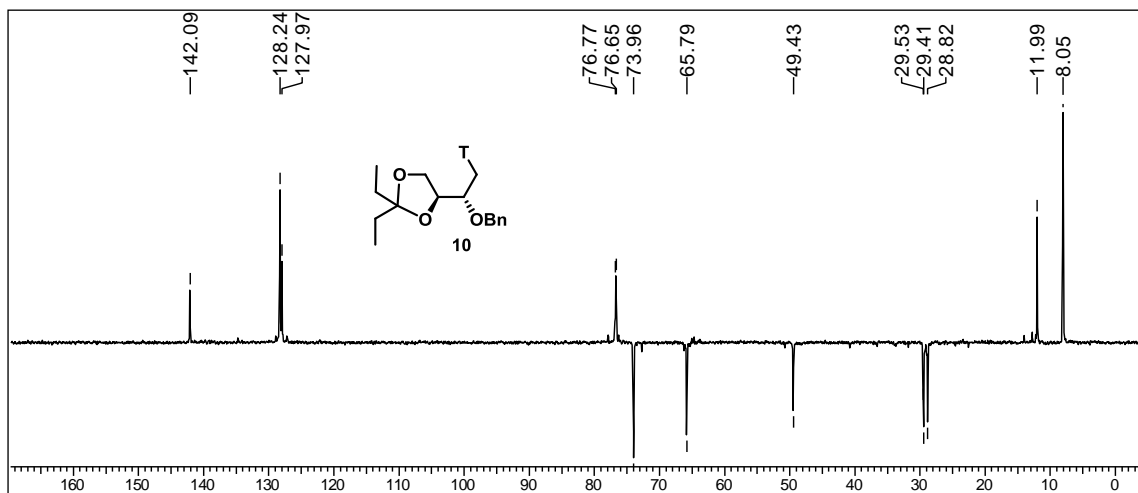
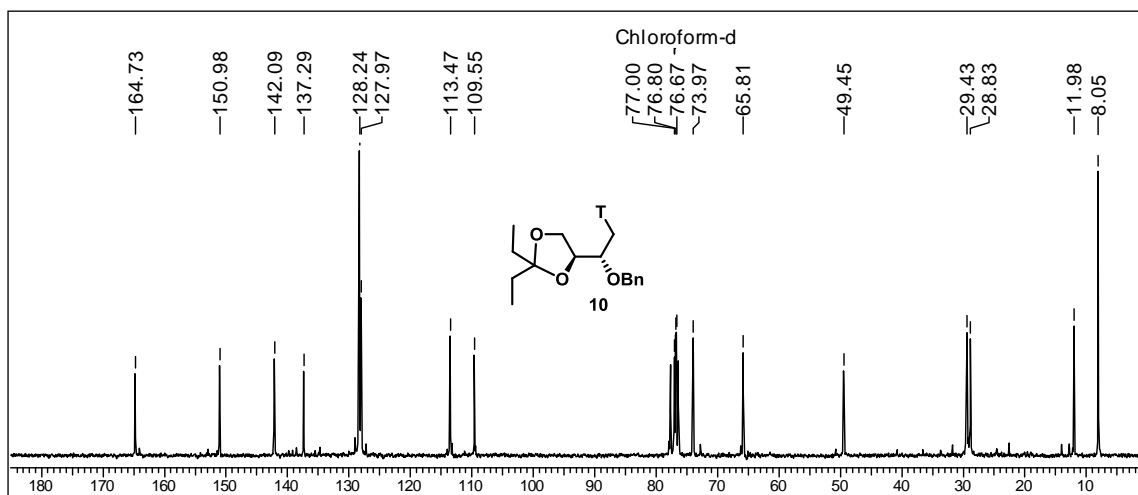
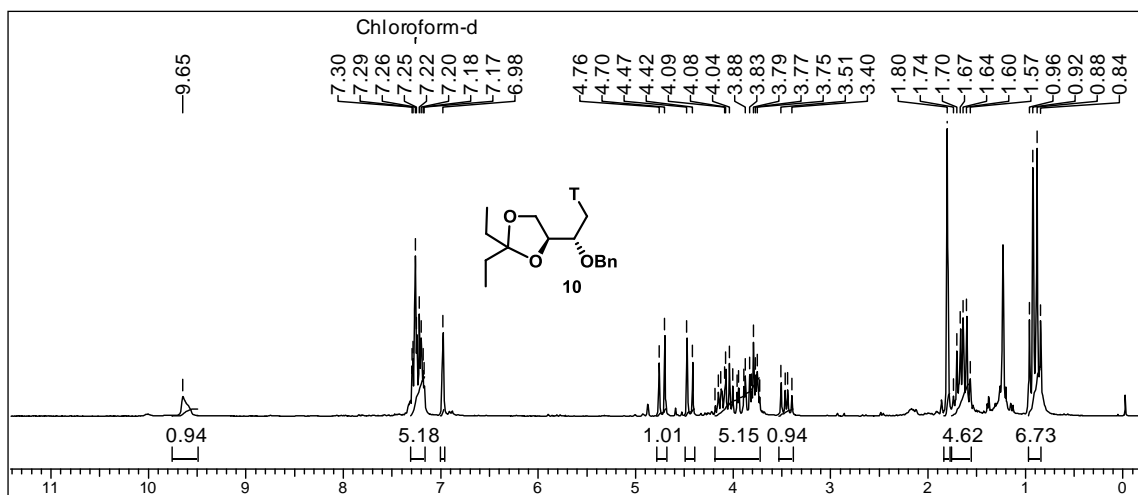
4.8 Appendix

Compounds - Spectral data	Page No.
9 - ^1H, ^{13}C NMR & DEPT	284
7 - ^1H, ^{13}C NMR & DEPT	285
10 - ^1H, ^{13}C NMR & DEPT	286
11 - ^1H, ^{13}C NMR & DEPT	287
12 - ^1H, ^{13}C NMR & DEPT	288
13 - ^1H, ^{13}C NMR & DEPT	289
14 - ^1H, ^{13}C NMR & DEPT	290
15 - ^{31}P	291
16 - ^1H, ^{13}C NMR & DEPT	292
17 - ^1H, ^{13}C NMR & DEPT	293
18 - ^1H, ^{13}C NMR & DEPT	294
19 - ^1H, ^{13}C NMR & DEPT	295
20 - ^1H, ^{13}C NMR & DEPT	296
10,11 - HRMS	297
12,13 - HRMS	298
14,15 - HRMS	299
16,18 - HRMS	300
19,20 - HRMS	301
ON1,ON2 - MALDI-TOF	302
ON3,ON4 - MALDI-TOF	303

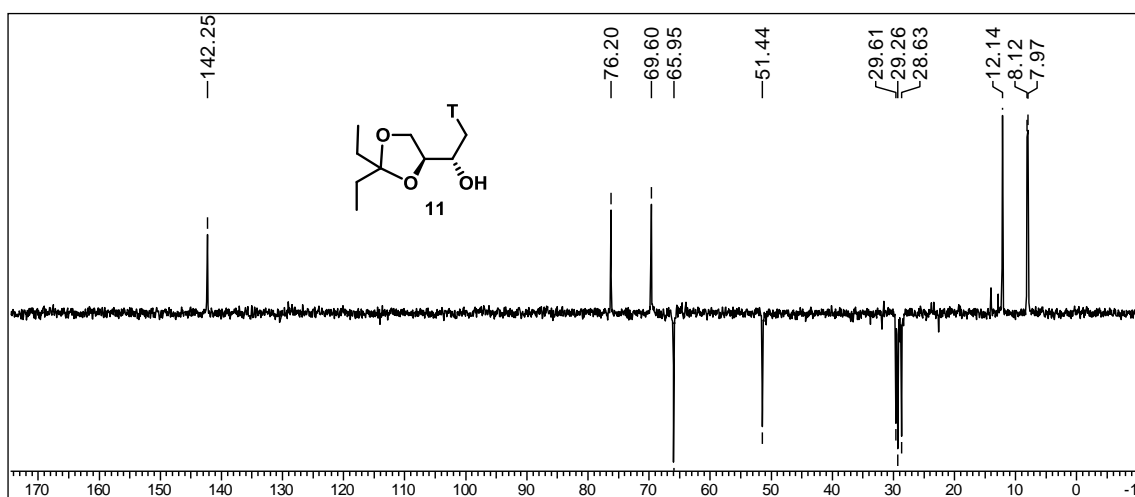
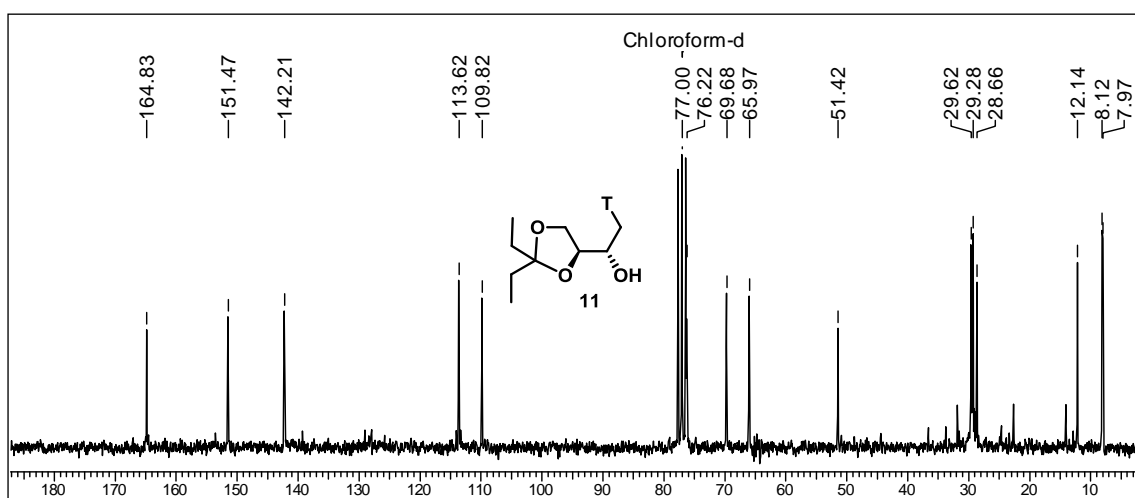
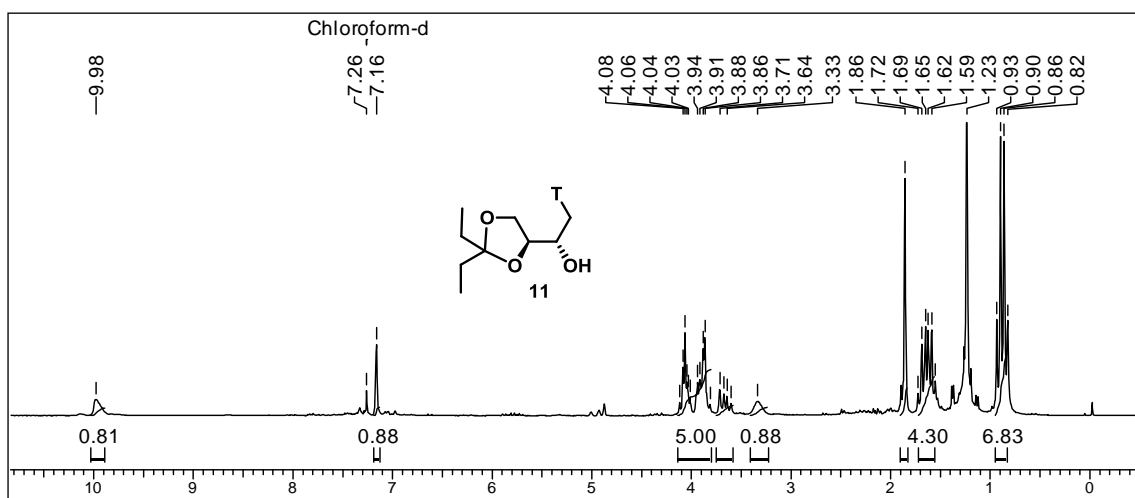
Chapter 4

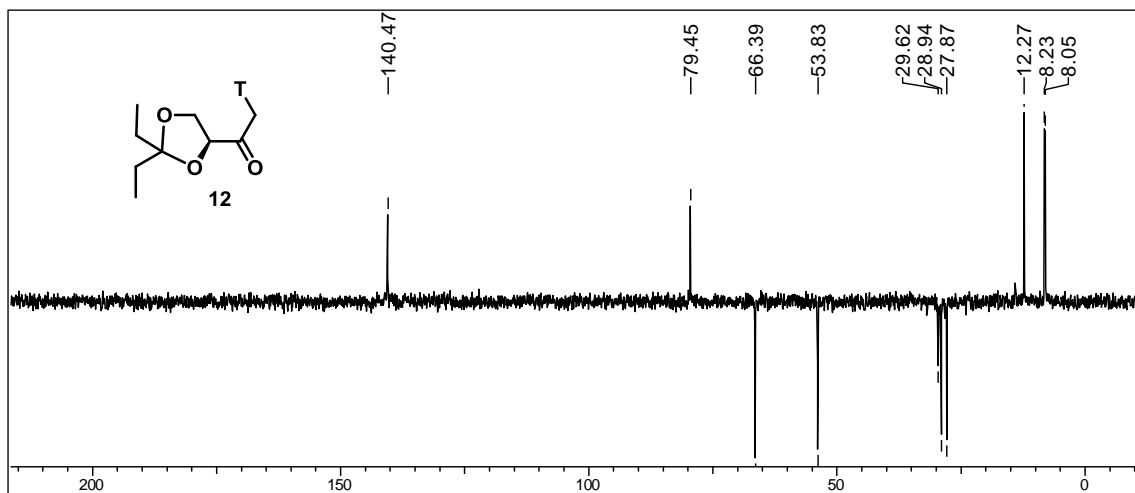
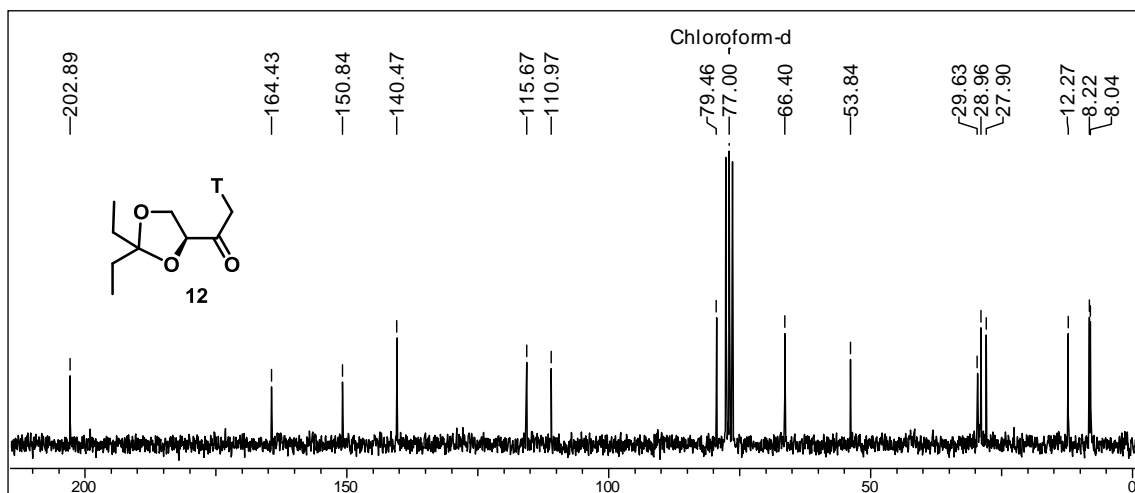
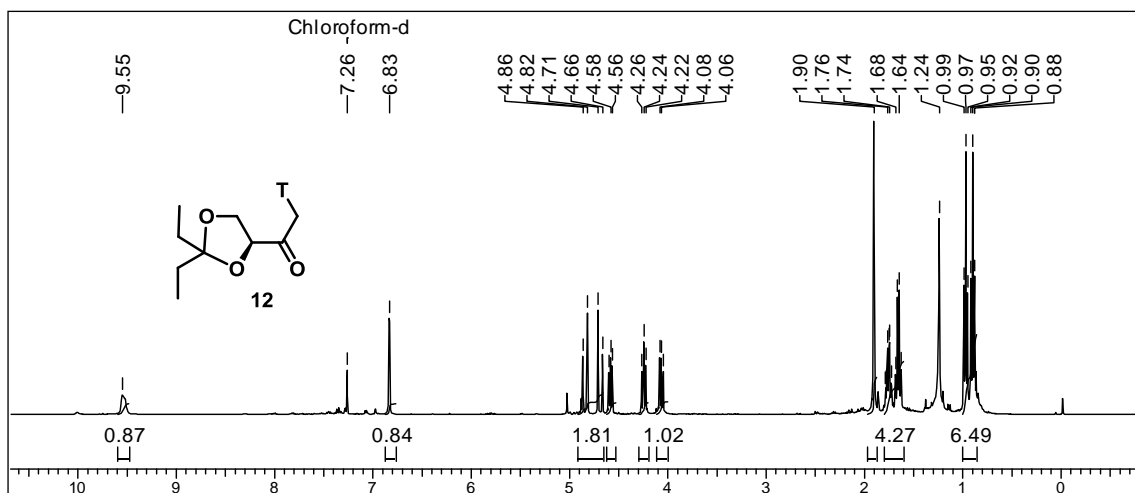


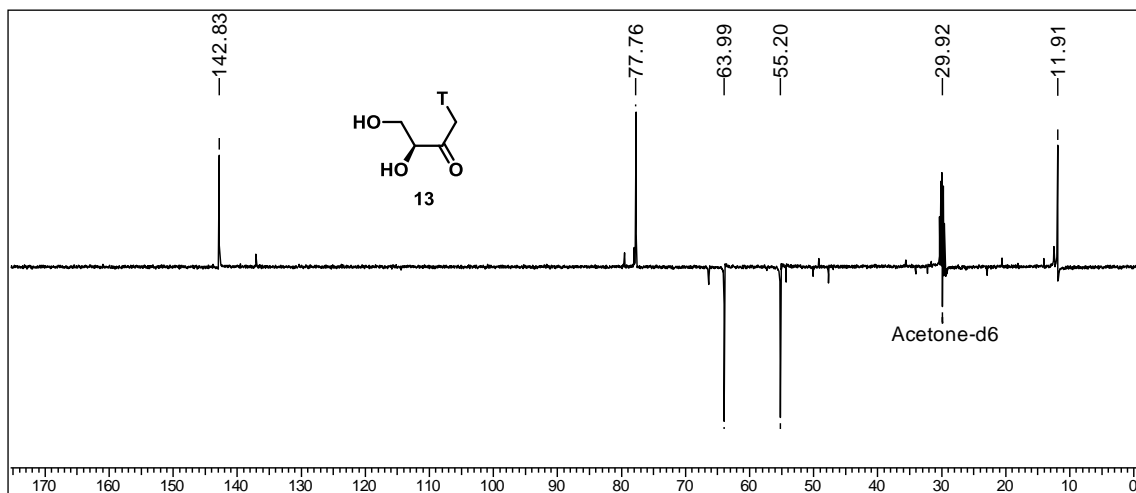
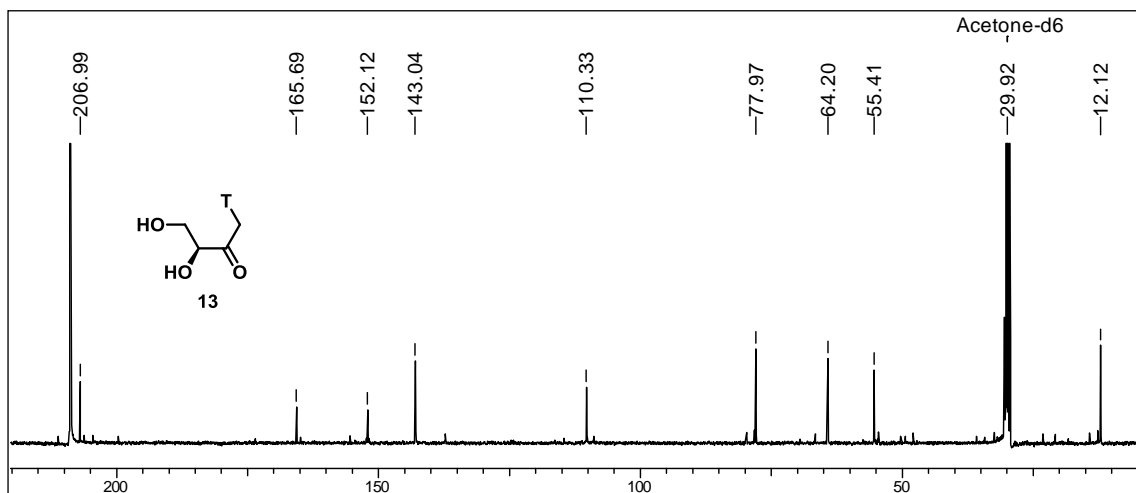
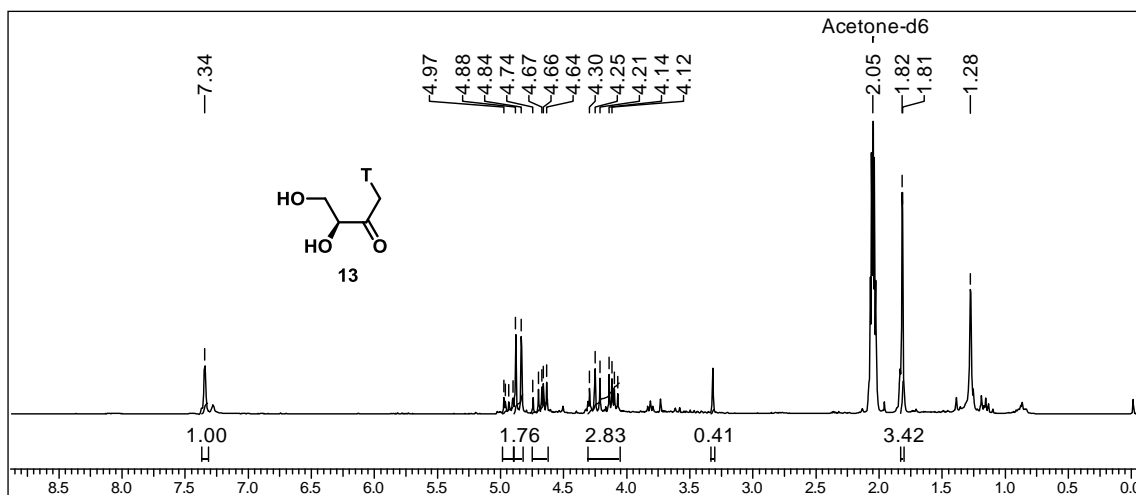


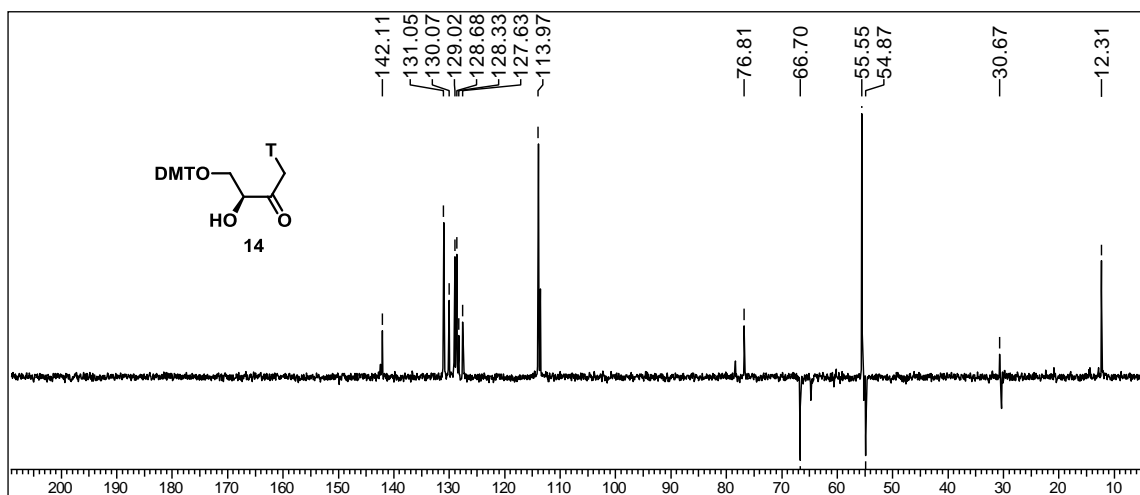
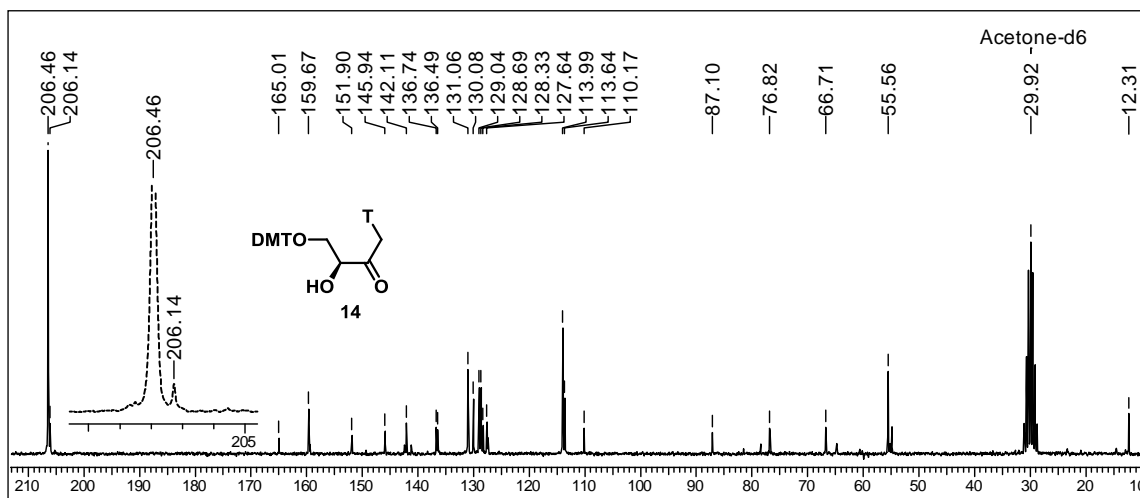
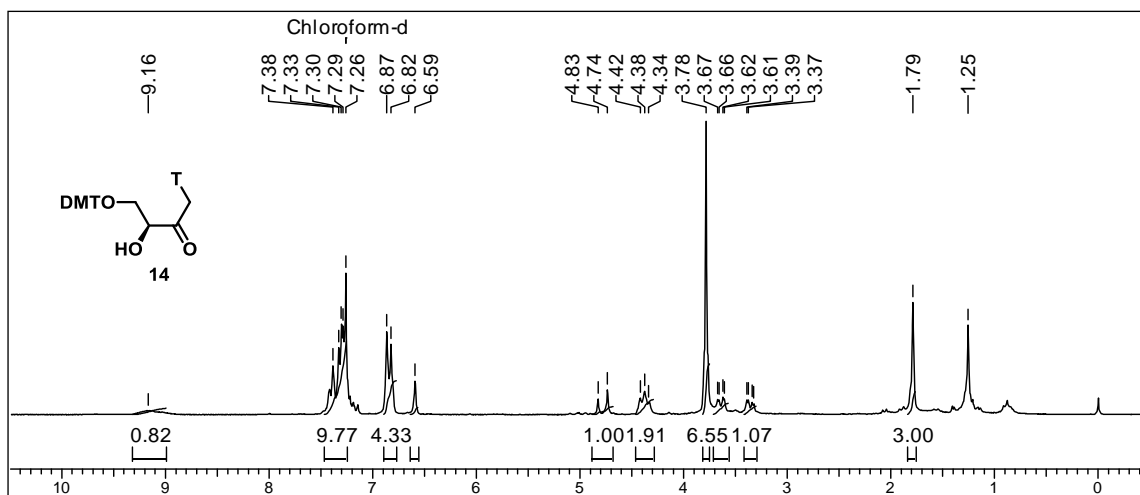


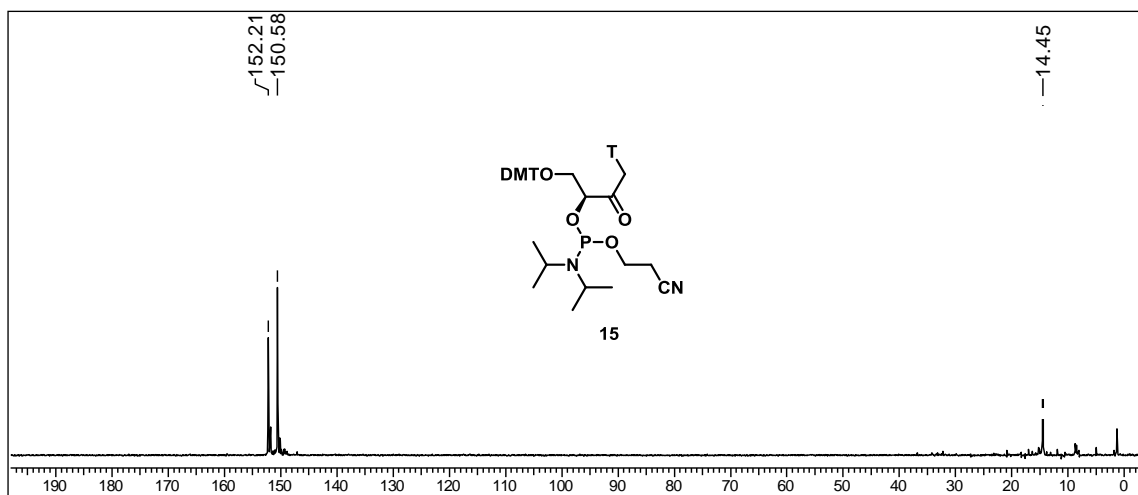
Chapter 4



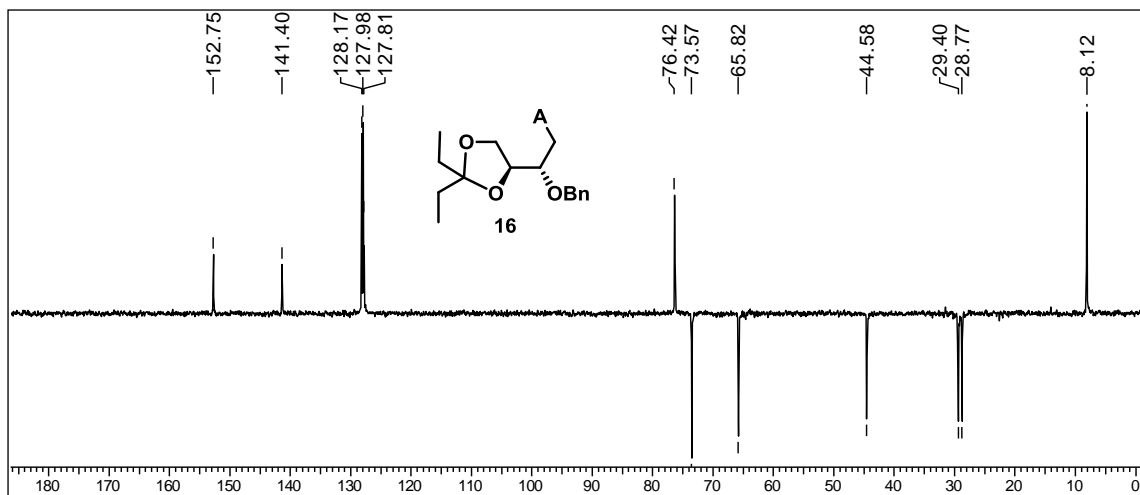
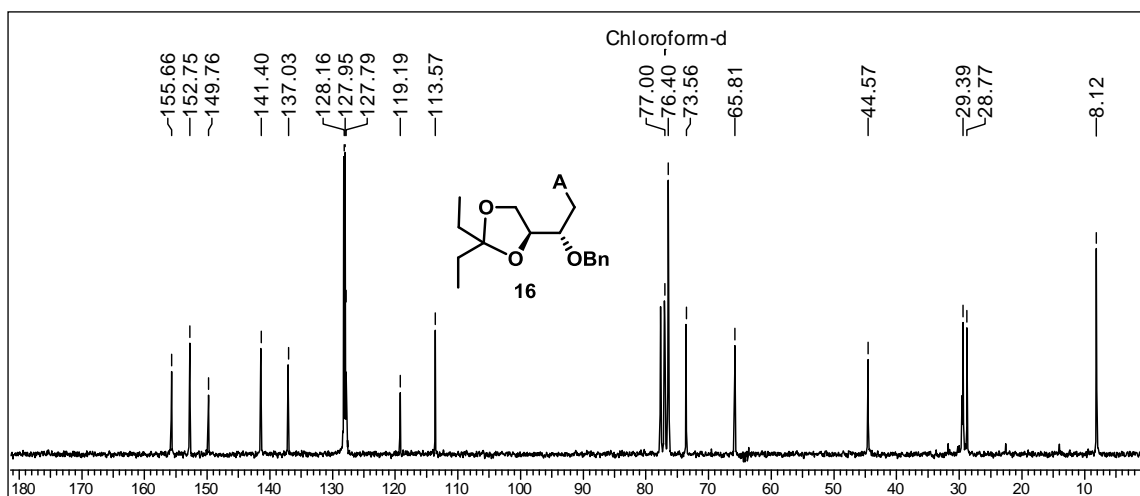
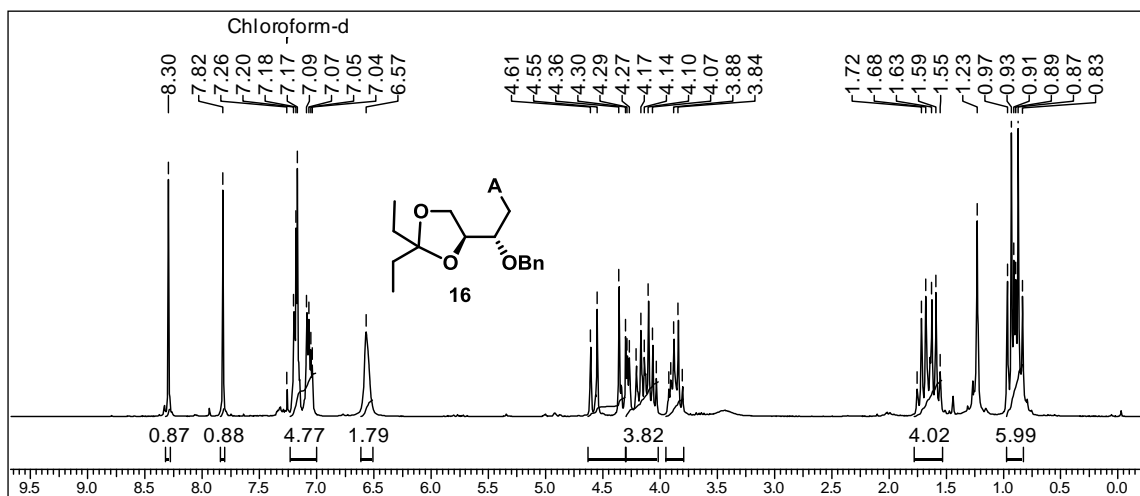


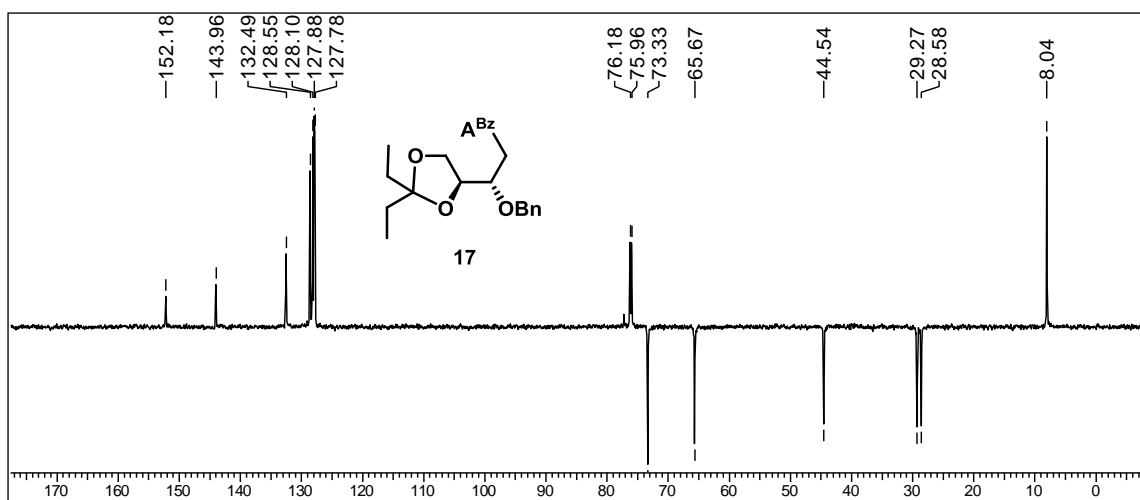
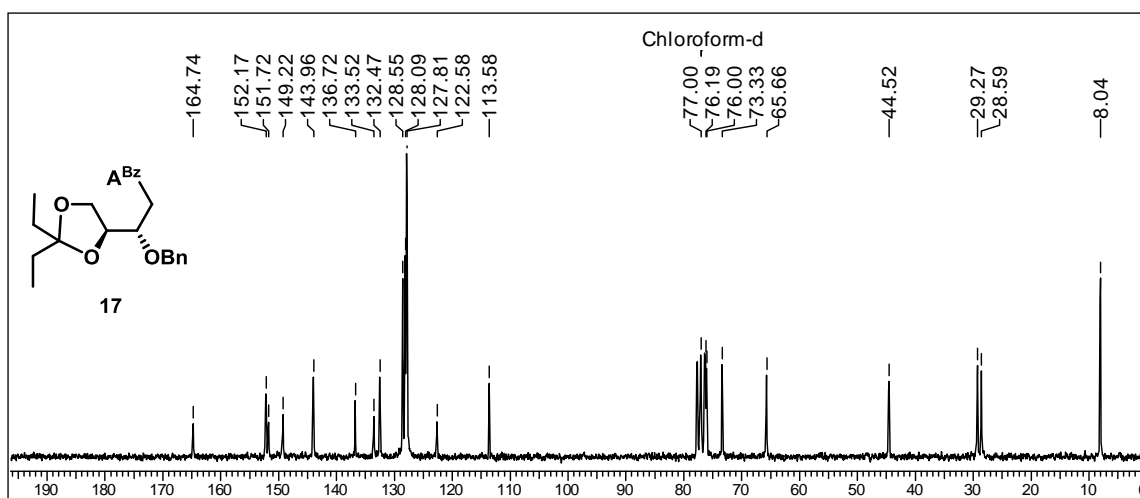
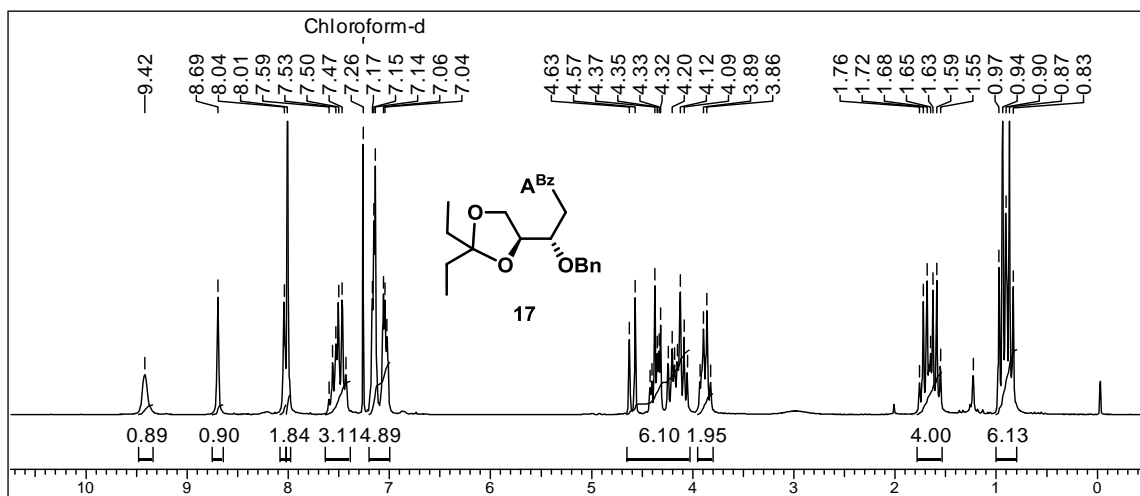


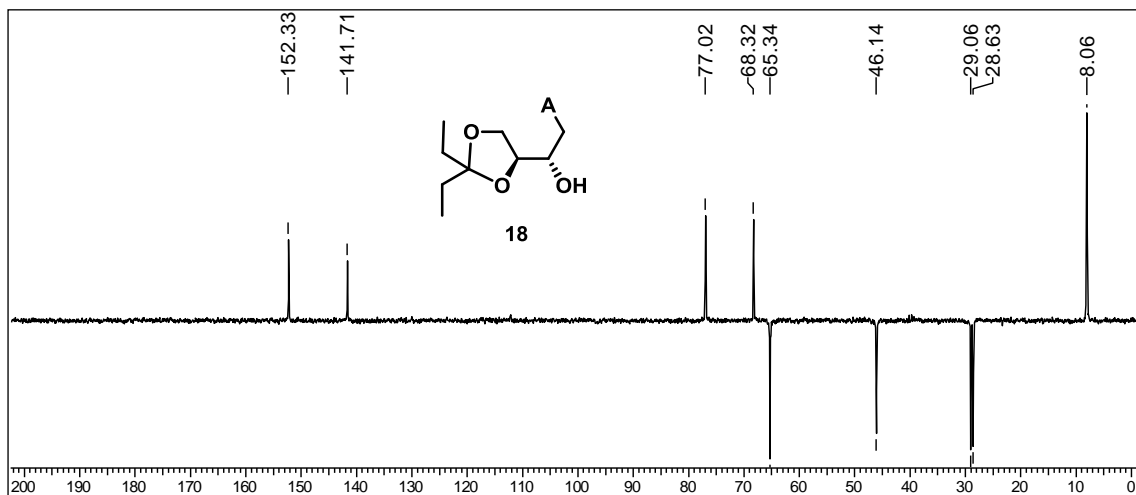
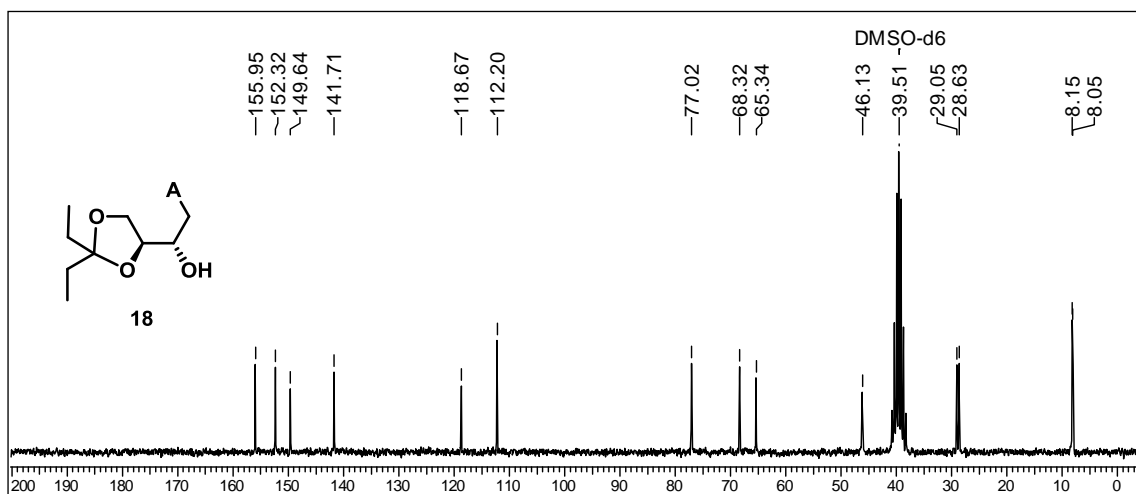
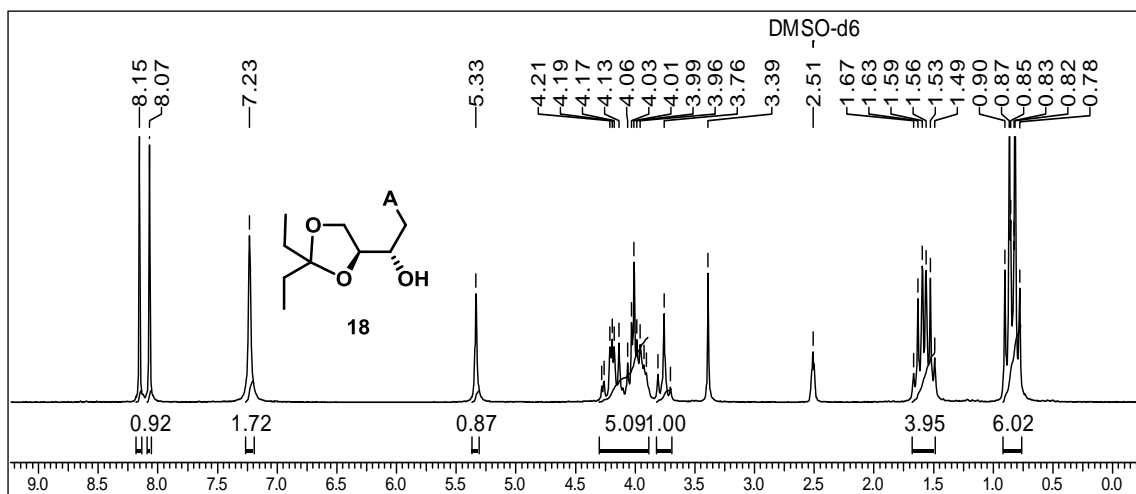


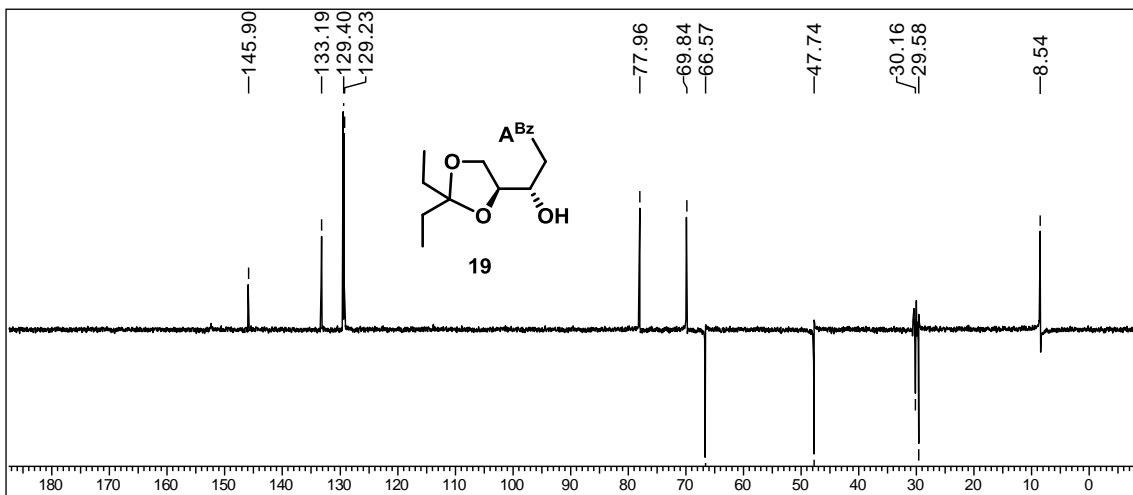
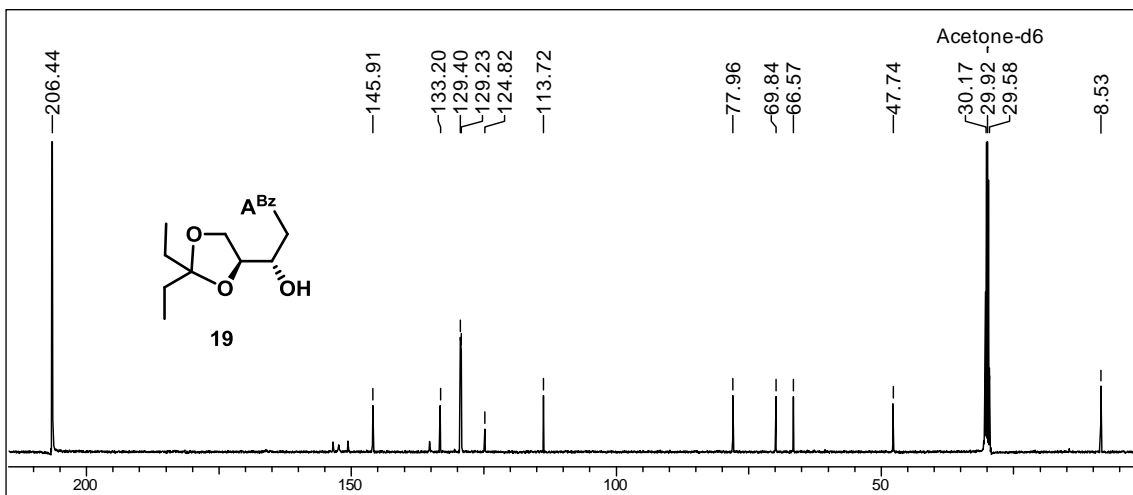
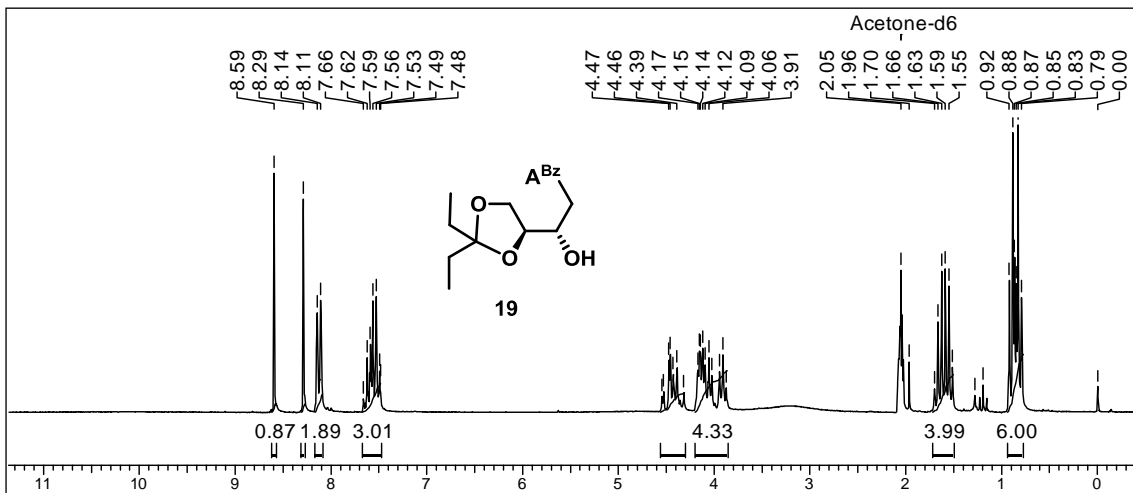


Chapter 4

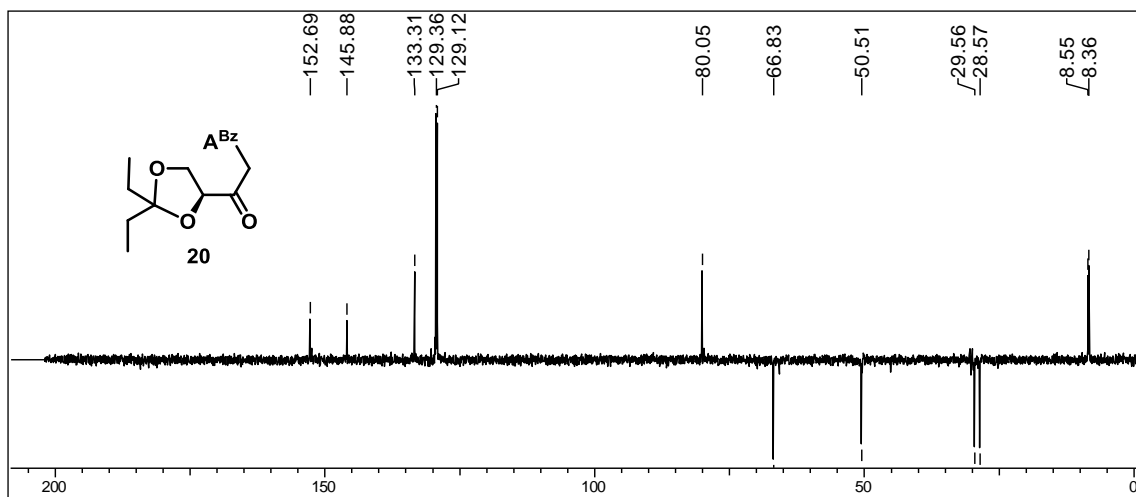
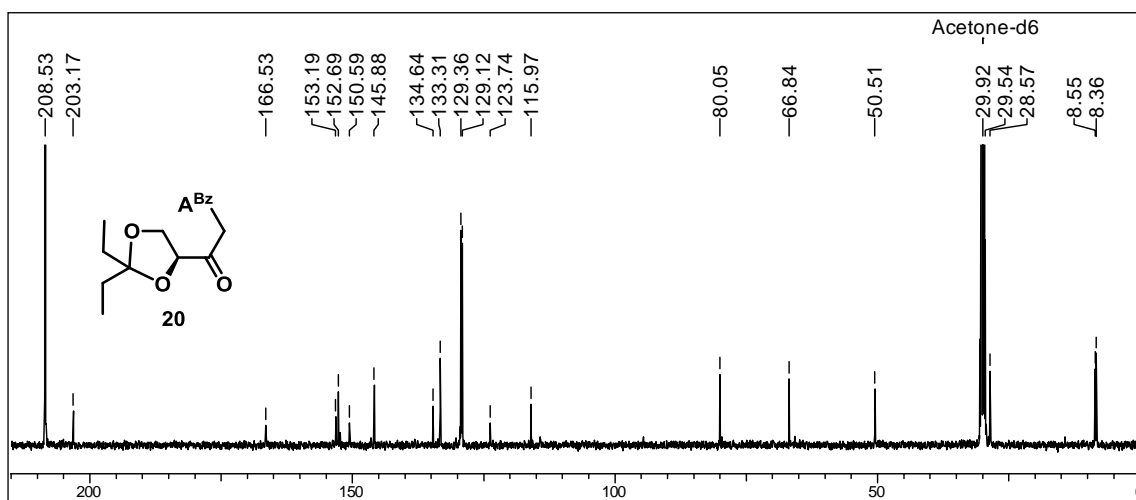
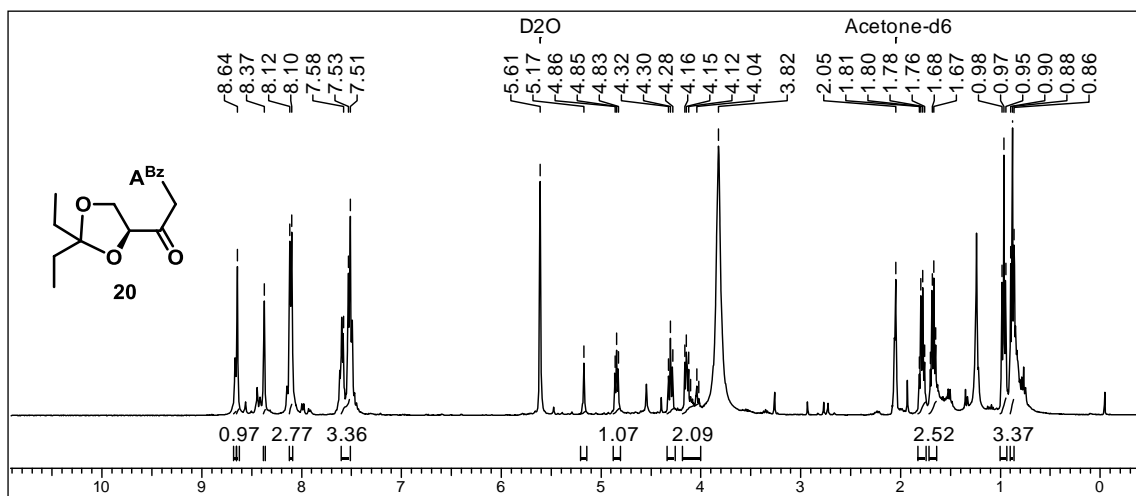




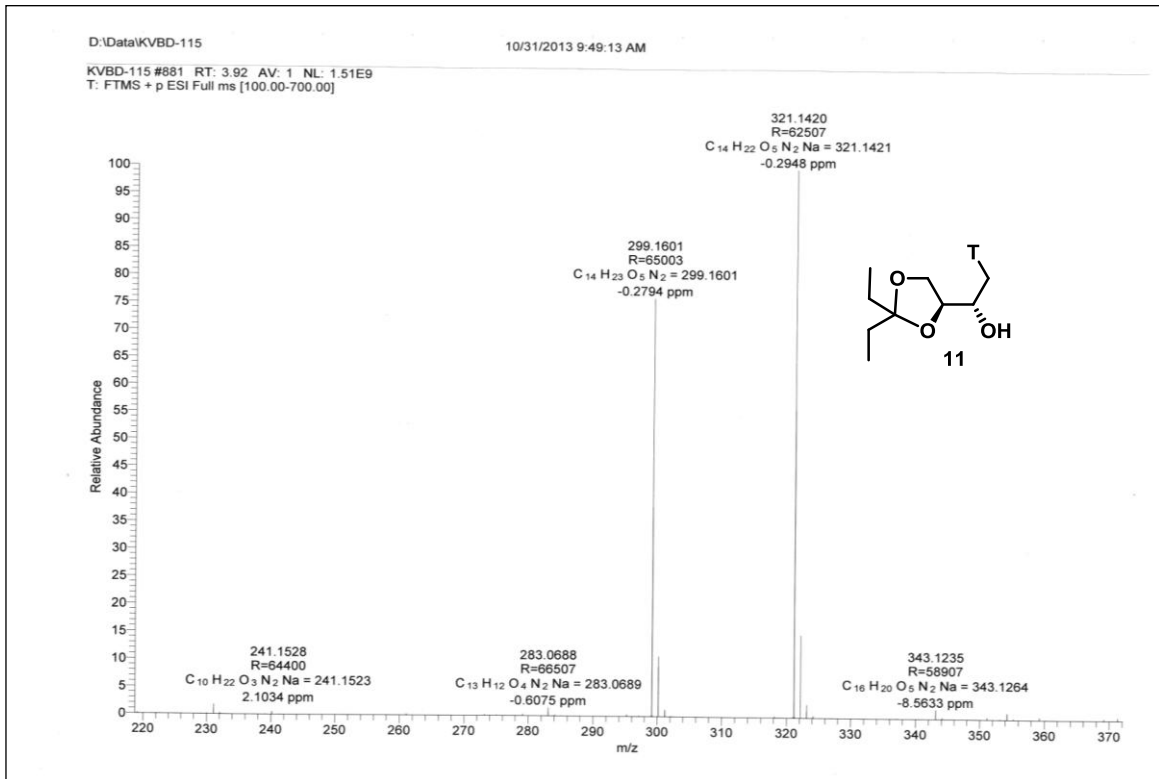
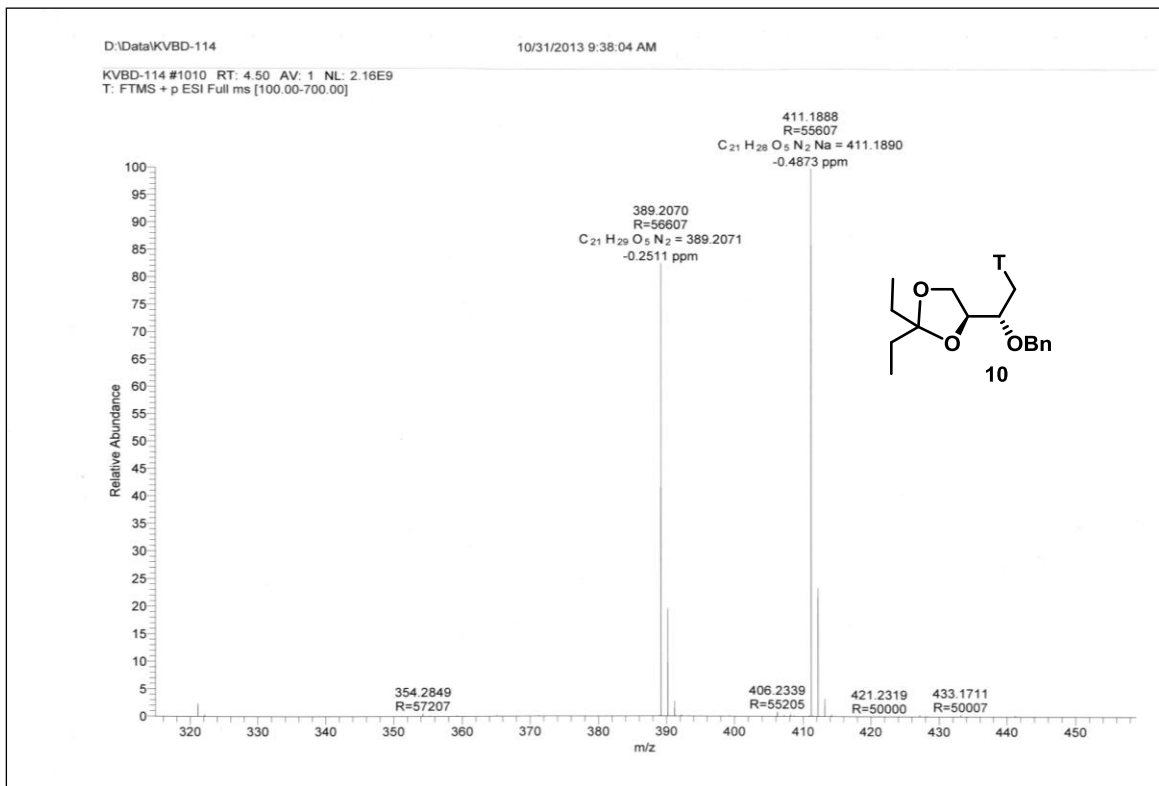




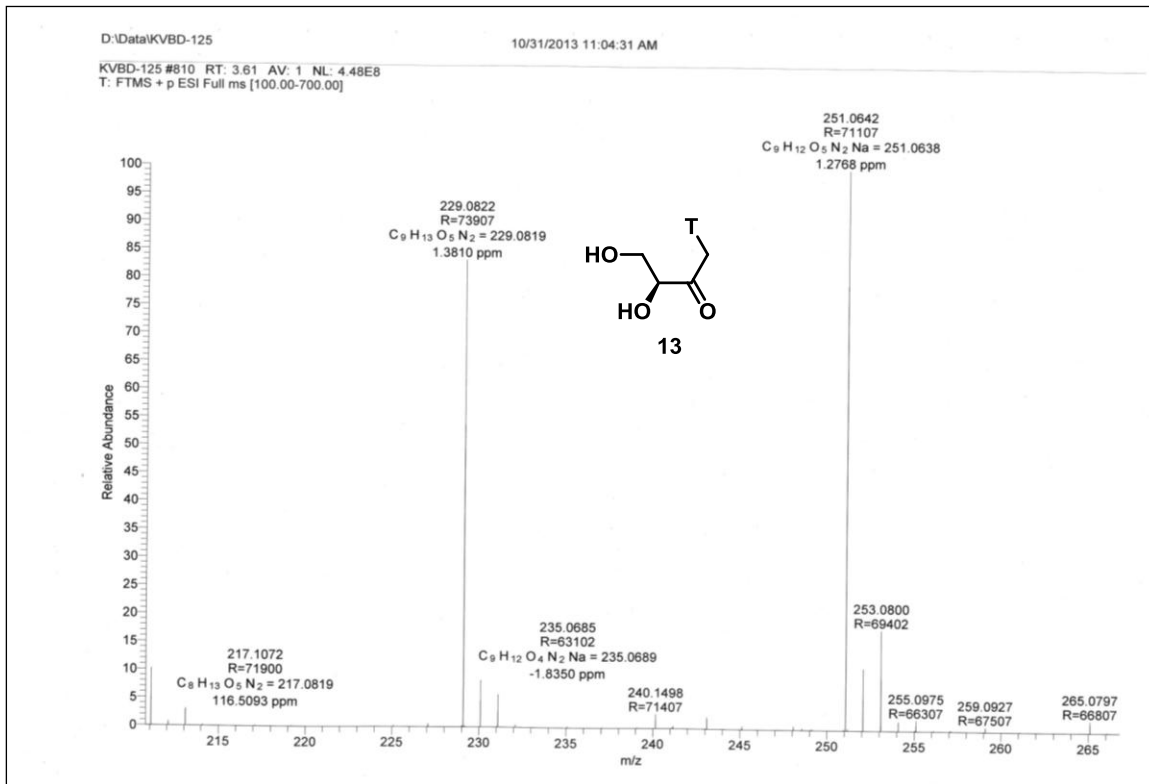
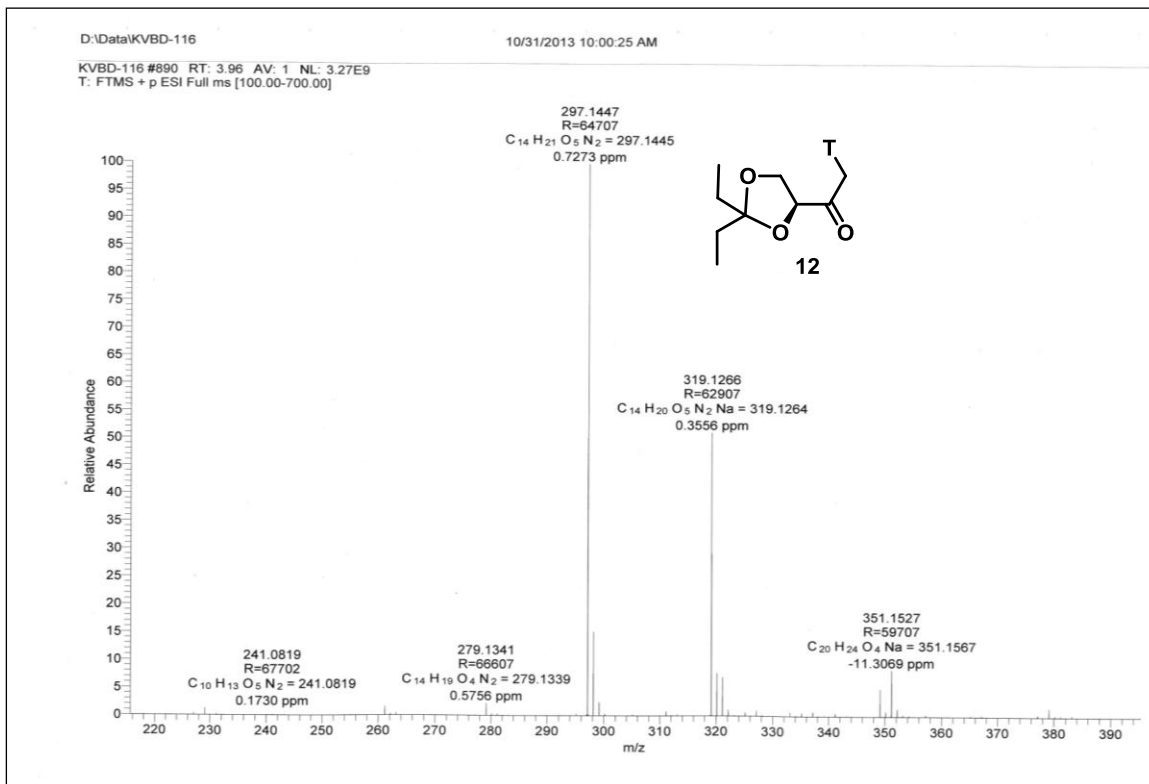
Chapter 4



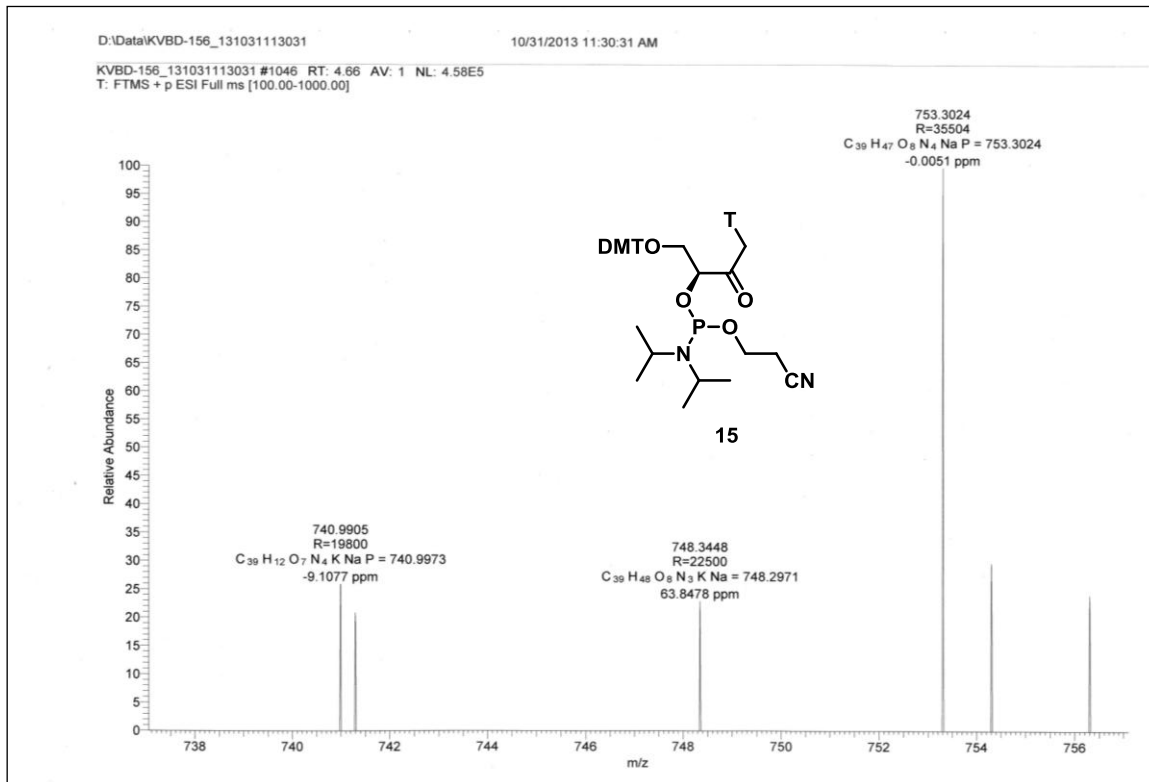
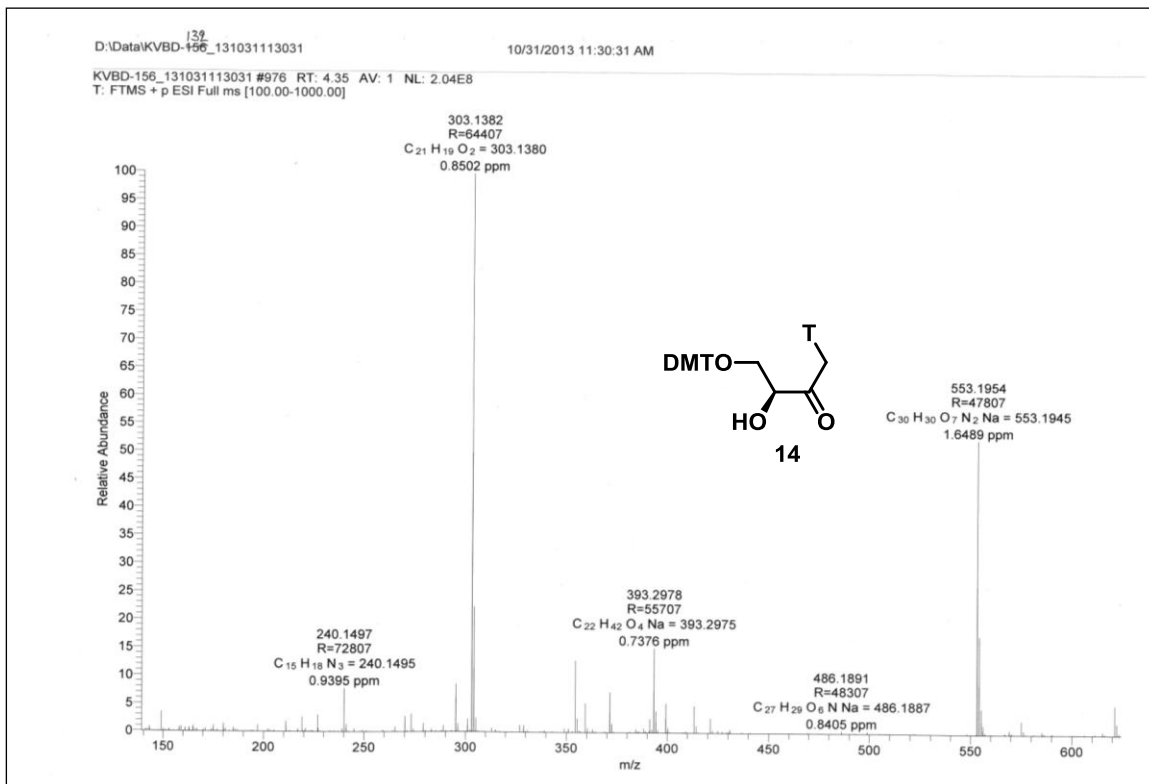
Chapter 4



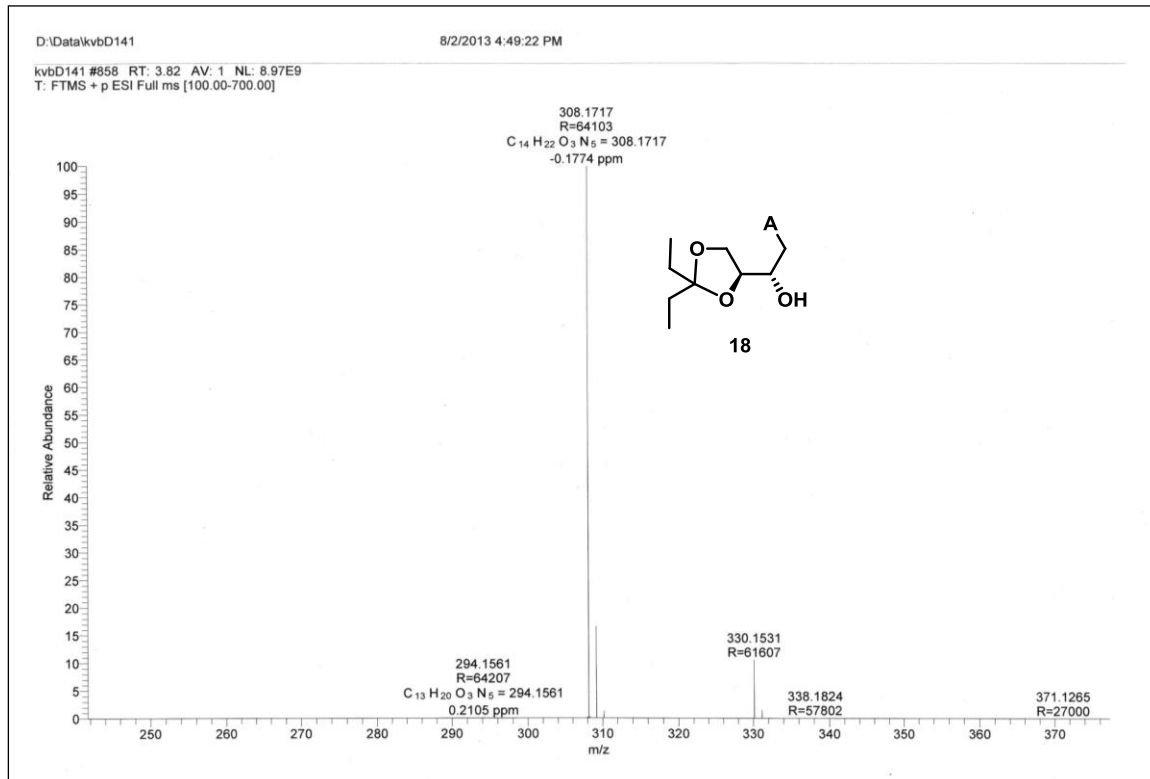
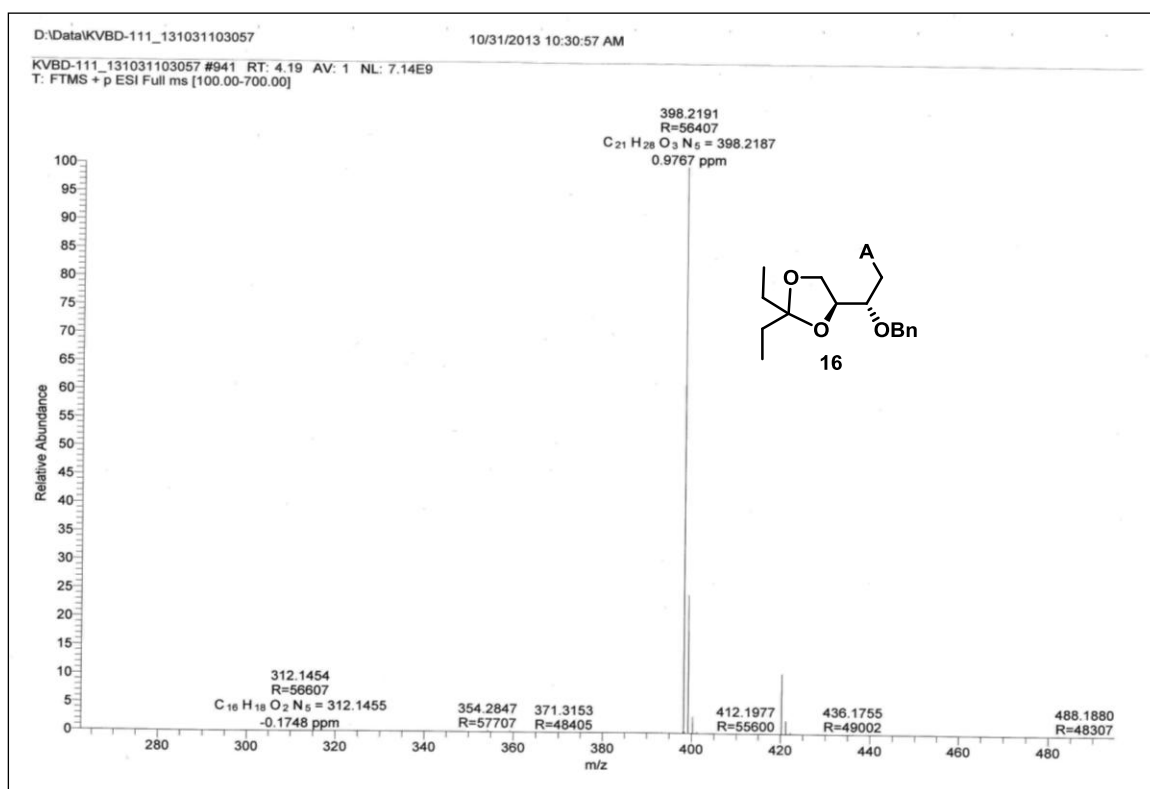
Chapter 4



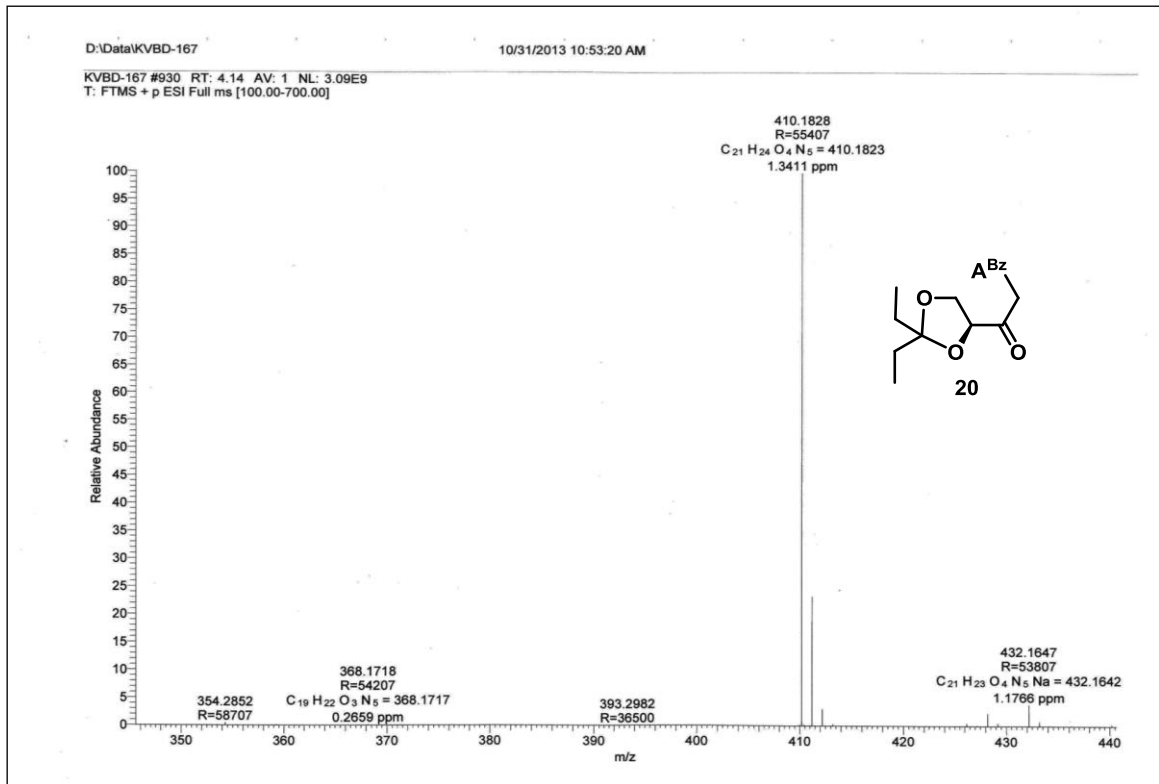
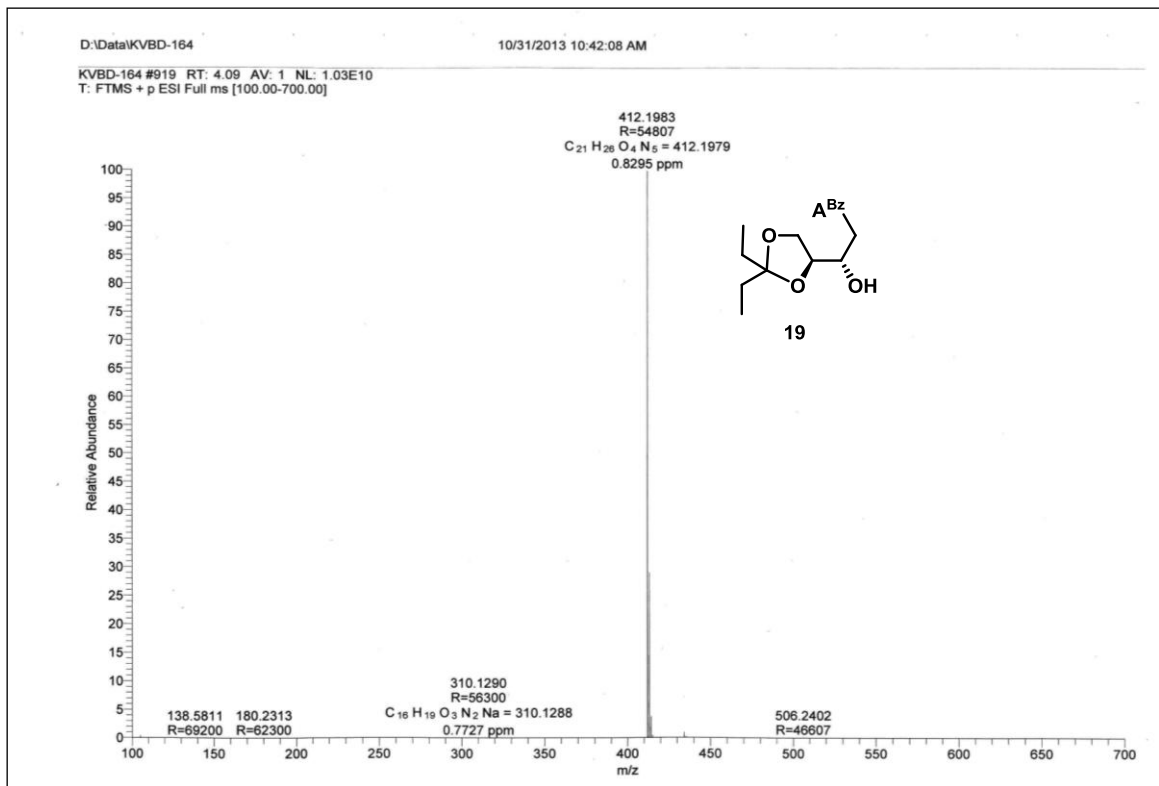
Chapter 4



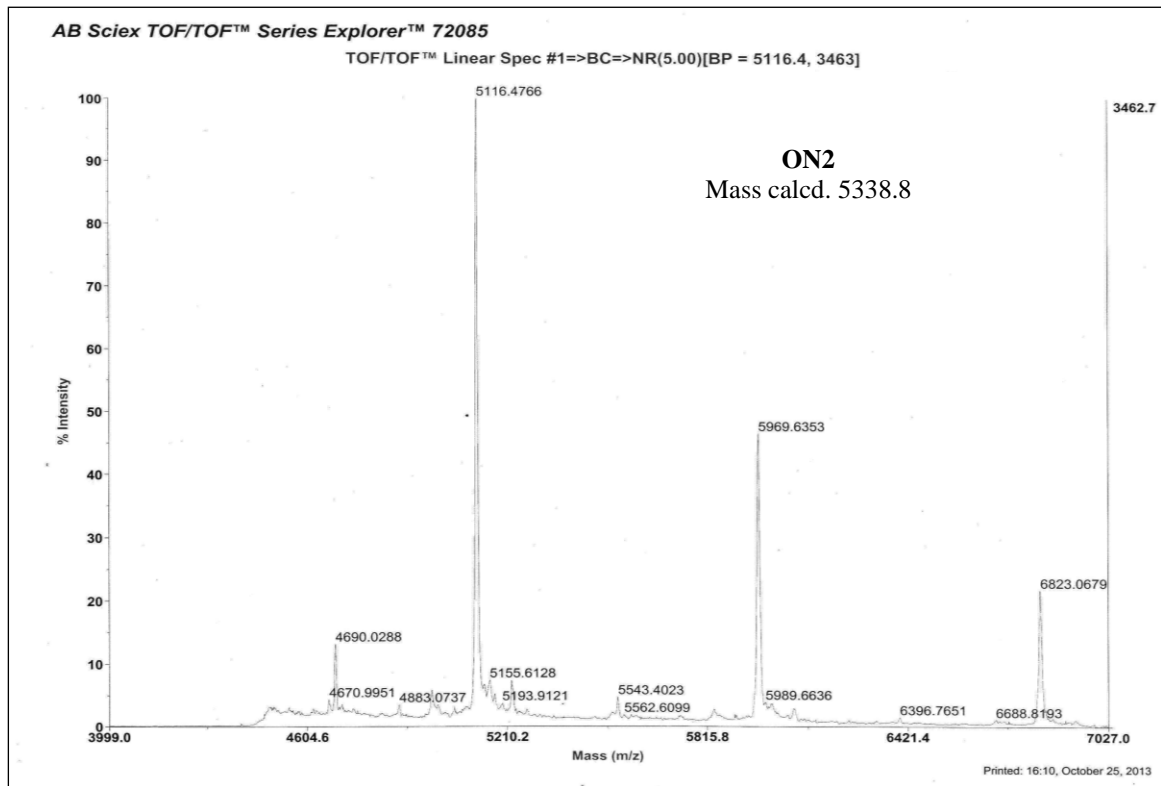
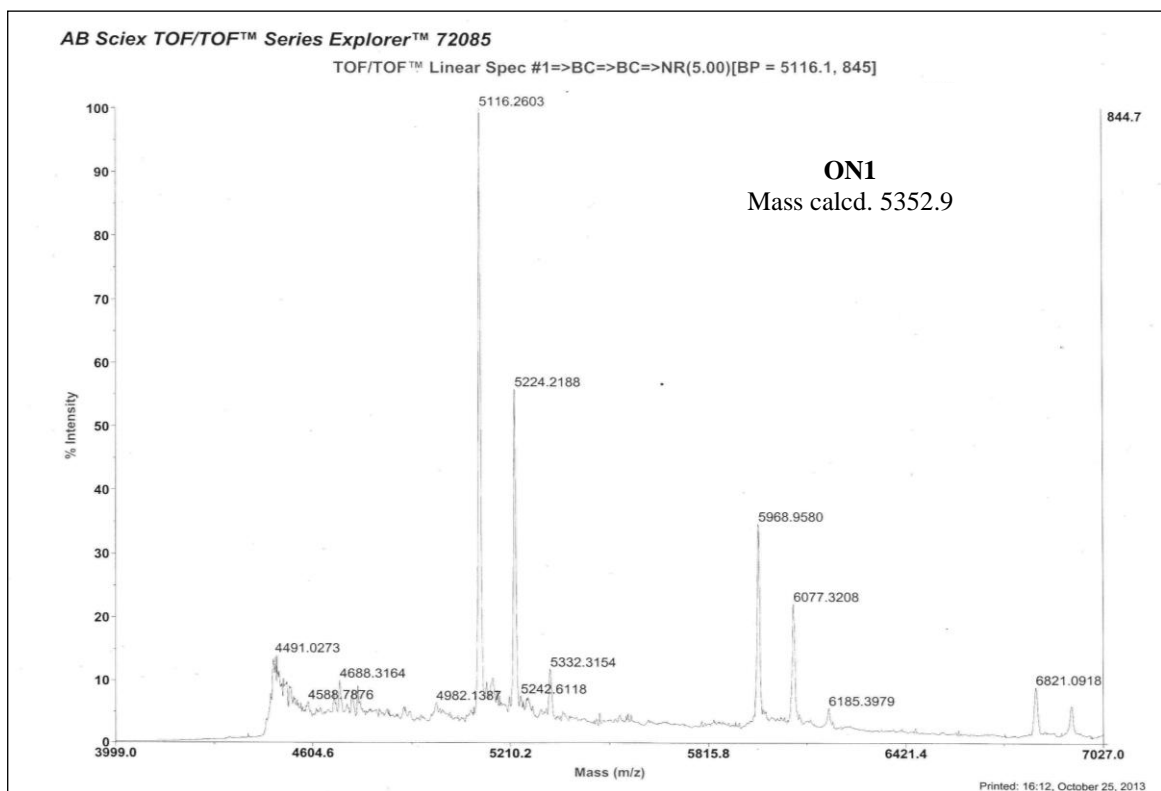
Chapter 4



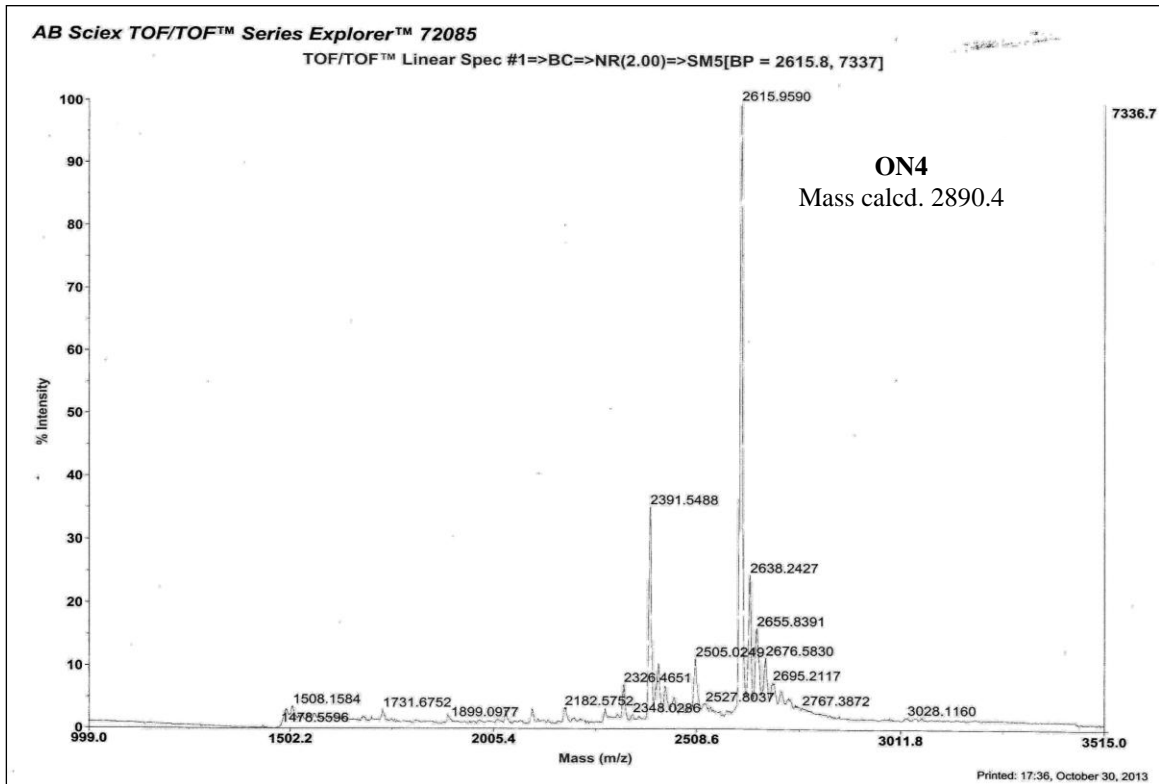
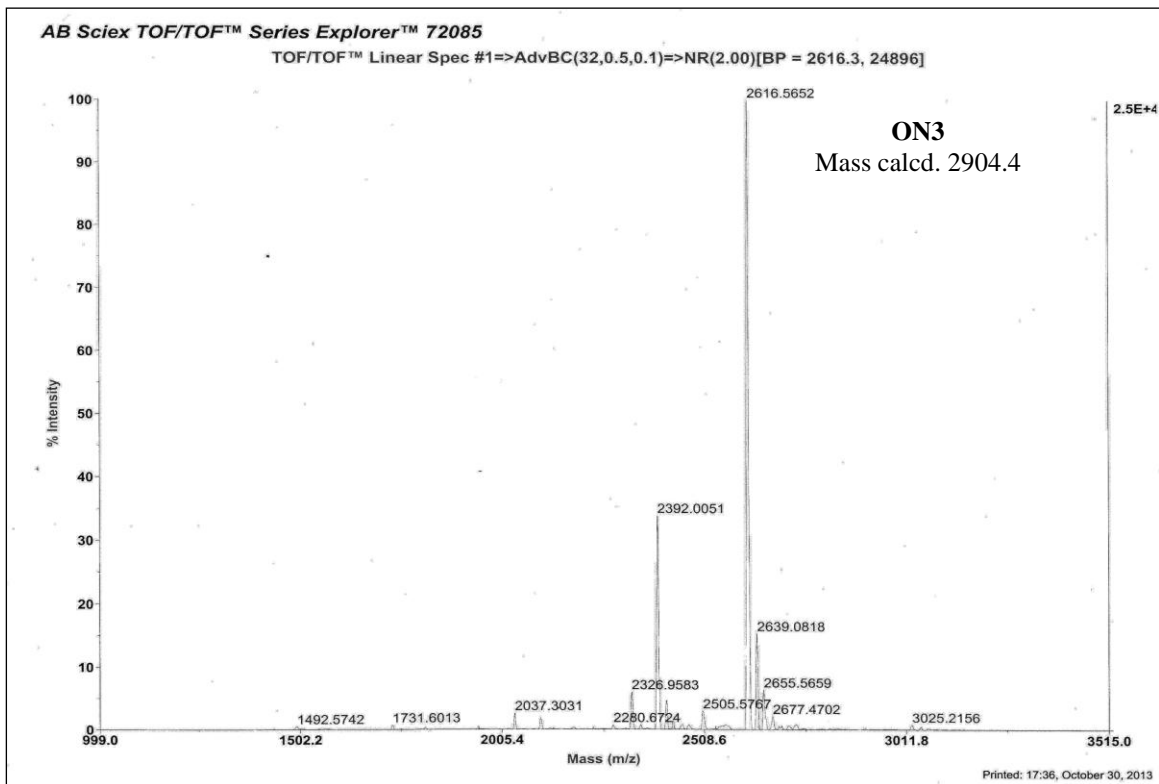
Chapter 4



Chapter 4



Chapter 4



4.10 References

1. (a)Woese, C., The evolution of the genetic code. *The genetic code (Eds.)* **1967**, Harper & Row, New York., 179–195. (b)Crick, F. H. C., *J. Mol. Biol.* **1968**, 38, 367-379. (c)Orgel, L. E., *J. Mol. Biol.* **1968**, 38, 381-393.
2. (a)Kruger, K.; Grabowski, P. J.; Zaug, A. J.; Sands, J.; Gottschling, D. E.; Cech, T. R., *Cell* **1982**, 31, 147-157. (b)Guerrier-Takada, C.; Gardiner, K.; Marsh, T.; Pace, N.; Altman, S., *Cell* **1983**, 35, 849-857.
3. Joyce, G. F., *Annu. Rev. Biochem.* **2004**, 73, 791-836.
4. (a)Engelhart, A. E.; Hud, N. V., *Cold Spring Harb. Perspect. Biol.* **2010**, 2:a002196. (b)Chaput, J. C.; Yu, H.; Zhang, S., *Chemistry & Biology* **2012**, 19, 1360-1371.
5. Joyce, G. F.; Schwartz, A. W.; Miller, S. L.; Orgel, L. E., *Proc. Natl. Acad. Sci. USA* **1987**, 84, 4398-4402.
6. Lazcano, A.; Miller, S. L., *Cell* **1996**, 85, 793-798.
7. Schoning, K.-U.; Scholz, P.; Guntha, S.; Wu, X.; Krishnamurthy, R.; Eschenmoser, A., *Science* **2000**, 290, 1347-1351.
8. Orgel, L., *Science* **2000**, 290, 1306-1307.
9. (a)Chaput, J. C.; Szostak, J. W., *J. Am. Chem. Soc.* **2003**, 125, 9274–9275. (b)Kempeneers, V.; Vastmans, K.; Rozenski, J.; Herdewijn, P., *Nucleic Acids Res.* **2003**, 31, 6221–6226.
10. Chaput, J. C.; Ichida, J. K.; Szostak, J. W., *J. Am. Chem. Soc.* **2003**, 125, 856–857.
11. (a)Ichida, J. K.; Horhota, A.; Zou, K. Y.; McLaughlin, L. W.; Szostak, J. W., *Nucleic Acids Res.* **2005**, 33, 5219–5225. (b)Horhota, A.; Zou, K.; Ichida, J. K.; Yu, B.; McLaughlin, L. W.; Szostak, J. W.; Chaput, J. C., *J. Am. Chem. Soc.* **2005**, 127, 7427–7434.
12. Yu, H.; Zhang, S.; Chaput, J. C., *Nat. Chem.* **2012**, 4, 183–187.
13. Brudno, Y.; Birnbaum, M. E.; Kleiner, R. E.; Liu, D. R., *Nat. Chem. Biol.* **2010**, 6, 148–155.
14. (a)Verheggen, I.; Van Aerschot, A.; Toppet, S.; Snoeck, R.; Janssen, G.; Balzarini, J.; De Clercq, E.; Herdewijn, P., *J. Med. Chem.* **1993**, 36, 2033–2040. (b)Verheggen, I.;

- Van Aerschot, A.; Van Meervelt, L.; Rozenski, J.; Wiebe, L.; Snoeck, R.; Andrei, G.; Balzarini, J.; Claes, P., *J. Med. Chem.* **1995**, 38, 826-835.
15. Krishnamurthy, R.; Pitsch, S.; Minton, M.; Miculka, C.; Windhab, N.; Eschenmoser, A., *Angew. Chem. Int. Ed.* **1996**, 35, 1537–1541.
16. (a)Hendrix, C.; Rosemeyer, H.; DeBouvere, B.; VanAerschot, A.; Seela, F.; Herdewijn, P., *Chemistry* **1997**, 3, 1513-1520. (b)Hendrix, C.; Rosemeyer, H.; Verheggen, I.; Seela, F.; VanAerschot, A.; Herdewijn, P., *Chemistry* **1997**, 3, 110-120.
17. Pinheiro, V. B.; Taylor, A. I.; Cozens, C.; Abramov, M.; Renders, M.; Zhang, S.; Chaput, J. C.; Wengel, J.; Peak-Chew, S. Y.; McLaughlin, S. H.; Herdewijn, P.; Holliger, P., *Science* **2012**, 336, 341-344.
18. Nelson, K. E.; Levy, M.; Miller, S. L., *Proc.Natl. Acad. Sci. U.S.A.* **2000**, 97, 3868-3871.
19. Banack, S. A.; Metcalf, J. S.; Jiang, L.; Craighead, D.; Ilag, L. L.; Cox, P. A., *PLoS One* **2012**, 7, e49043.
20. Ueda, N.; Kawabata, T.; Takemoto, K., *J. Heterocycl. Chem.* **1971**, 8, 827-829.
21. Zhang, L.; Peritz, A.; Meggers, E., *J. Am. Chem. Soc.* **2005**, 127, 4174-4175.
22. Schlegel, M. K.; Peritz, A. E.; Kittigowittana, K.; Zhang, L.; Meggers, E., *ChemBioChem* **2007**, 8, 927-932.
23. Chen, J. J.; Cai, X.; Szostak, J. W., *J. Am. Chem. Soc.* **2009**, 131, 2119-2121.